

BIRLA CENTRAL LIBRARY
PILANI [RAJASTHAN]

Class No 530.72

Book No V15414

Accession No. 38074

**INTRODUCTION TO
THEORETICAL
AND
EXPERIMENTAL
OPTICS**

INTRODUCTION TO
THEORETICAL
AND
EXPERIMENTAL
OPTICS

BY

Joseph Valasek

*Professor of Physics
University of Minnesota*

John Wiley & Sons, Inc., New York
Chapman & Hall, Limited, London

COPYRIGHT, 1949
BY
JOHN WILEY & SONS, INC.

All Rights Reserved

*This book or any part thereof must not
be reproduced in any form without
the written permission of the publisher.*

PRINTED IN THE UNITED STATES OF AMERICA

PREFACE

This textbook is intended primarily for students who have completed at least a course in general physics and in calculus. It includes the principles of geometrical optics and physical optics, as well as experiments in optics and spectroscopy. I have found such a combination to be of more practical usefulness to the student of physics than a purely theoretical treatise or the usual type of laboratory manual. The mathematical treatment is generally of an elementary nature except where higher mathematics is distinctly advantageous or indispensable. An attempt has been made to discuss the various topics in a manner which will be most understandable and useful to an undergraduate at the advanced physics level and to the first-year graduate student. In general, graduate students in physics should find much material here which they may study with profit, since some of this is often omitted from advanced courses in theoretical physics.

This book may be used as a manual in a course which consists primarily of laboratory work. In this case, the student will find in the theoretical portion a justification of the experimental procedures and derivations of the formulas used. It seems desirable to have this material available for convenient reference even in a laboratory course. Furthermore, supplementary or alternative experiments may be devised more readily since the theoretical background may be found in the text.

It is also intended that many parts of the text may serve as an introduction to applied optics or to optical engineering. This field is still relatively unrecognized as a specialty as compared with those branches of engineering which deal with the applications of mechanics, heat, electricity, and acoustics. Nevertheless, there are so many applications of optics in industry and in the research laboratory that an engineer, physicist, or research scientist in general will often find a foundation in this subject extremely useful. Even in such indirectly related fields as electron optics, microwaves, and television, the principles of optics find many practical applications.

I have been teaching courses in the various branches of optics at elementary and advanced levels during the past twenty-nine years and, at one time or other, have used most of the standard textbooks and many references to supplementary material. This book is written primarily for the purpose of presenting in a single volume what appears to be the most effective treatment of the various optical topics on an intermediate level. In a course which follows directly after the usual type of course in general physics, some of the mathematical developments may have to be omitted. This may be accomplished by simply summarizing and discussing the results without formal proof. Although most of the material presented here is classical, there is frequent discussion of modern developments, as far as seems suitable in a book of this scope. The theory is usually developed from first principles with the exception of the optical applications of electrodynamics and quantum mechanics. In the first case, Maxwell's equations are taken as the foundation for the development of a theory of crystal optics, Fresnel's equations for reflection, and the optical properties of metallic media. On the other hand, quantum-theoretical results are generally quoted without proof. Since I am particularly interested in x-rays, photographic optics, and ophthalmic lenses, these subjects are treated in somewhat greater detail than is customary in texts dealing with optics.

Specialists in the various subdivisions of optics will no doubt find their specialties treated too briefly or on too elementary a plane. This can hardly be avoided in a book of this size which aims to serve as an introduction to many branches of optics. Those who require a more extensive treatment of particular subjects are referred to the bibliography for supplementary material.

Portions of the material on crystal optics and interference in polarized light first appeared in an article by the author in the *Journal of Applied Physics*, **10**, pages 209–221 (1939). I take this opportunity to thank Dean Elmer Hutchisson, the editor, for permission to use this material here. I also wish to acknowledge the assistance of my wife, Leila E. Valasek, who read the proof and checked many of the formulas.

JOSEPH VALASEK

University of Minnesota
September 16, 1949

CONTENTS

I. GEOMETRICAL OPTICS

CHAPTER 1. FUNDAMENTAL PRINCIPLES	1
Postulates and Basic Laws . . . Fermat's Principle . . . Reflection and Refraction . . . Absorption	
CHAPTER 2. IDEAL OPTICAL SYSTEMS	10
Images . . . Collinear Transformation . . . Principal Foci and Principal Planes . . . Angular Magnification and Nodal Points . . . Classification of Optical Systems	
CHAPTER 3. COMPUTATION OF CONSTANTS	20
Trigonometrical Ray Tracing . . . First Order Theory . . . Combination of Optical Systems . . . The Simple Lens in Air . . . Dioptric Power of a Lens . . . Alternative Computation of Focal Length . . . Telescopic System	
CHAPTER 4. PHYSICAL LIMITATIONS OF OPTICAL IMAGES	36
General Considerations . . . Abbe's Sine Condition . . . Aberrations . . . Spherical Aberration . . . Coma . . . Distortion . . . Astigmatism . . . Curvature of Image . . . Chromatic Aberration	
CHAPTER 5. APERTURES OF OPTICAL SYSTEMS	53
Diaphragms or Stops . . . Entrance Pupil and Exit Pupil . . . Field of View . . . Depth of Field and Focus . . . Perspective	
CHAPTER 6. PHOTOMETRY AND IMAGE BRIGHTNESS	61
Photometric Quantities . . . Brightness of Images . . . Illuminance of Images . . . Normal Brightness . . . Normal Magnification	
CHAPTER 7. OPTICAL INSTRUMENTS	73
The Eye . . . Spectacle Lenses . . . The Simple Magnifier . . . Eyepieces . . . The Compound Microscope . . . Telescopes . . . Reflecting Telescopes . . . Prism Spectroscopes and Spectrographs	
CHAPTER 8. STEREOSCOPY	101
Binocular Vision . . . Stereoscopic Camera and Viewer . . . Vectographs	

II. PHYSICAL OPTICS

CHAPTER 9. WAVE PROPAGATION	106
Nature of Light . . . Wave Theory . . . Wave and Group Velocity . . . Measurement of the Speed of Light	
CHAPTER 10. INTERFERENCE OF LIGHT	115
Young's Experiment . . . Addition of Waves . . . Fresnel's Mirrors and Biprism . . . Michelson Stellar Interferometer . . . Rayleigh Gas Interferometer . . . Interference in Films and Plates . . . Localization of Fringes . . . Applications of Fringes of Constant Thickness . . . Interference with Large Path Difference . . . The Michelson-Morley Experiment . . . Visibility of the Fringes . . . Widths of Spectral Lines . . . Jamin Interferometer . . . The Fabry-Perot Interferometer	
CHAPTER 11. DIFFRACTION	156
Diffraction Phenomena . . . Single Slit in Parallel Light . . . Circular Aperture . . . Resolving Power . . . The Double Slit . . . Diffraction Grating . . . Resolving Power of a Grating . . . Fresnel Diffraction Phenomena . . . Kirchhoff's Formulation of Fresnel's Theory . . . Diffraction of Microwaves . . . Diffraction of X-rays by Crystals	
CHAPTER 12. POLARIZATION AND DOUBLE REFRACTION	193
Polarized Light . . . Electromagnetic Theory . . . Reflection and Refraction . . . Doubly Refracting Media . . . Crystal Optics . . . Ray Velocity Surface . . . Index Surface . . . Velocities inside Aniso- tropic Media . . . Double Refraction . . . Plane Polariscopes . . . Elliptically and Circularly Polarized Light . . . Babinet Compensator . . . Interference in Circularly Polarized Light	
CHAPTER 13. OPTICAL ACTIVITY	227
Optically Active Materials . . . Polarimetry	
CHAPTER 14. REFRACTION AND DISPERSION	232
Theory of Refraction . . . Refractivity . . . Dispersion . . . Quantum Theory of Dispersion . . . X-ray Dispersion	
CHAPTER 15. OPTICAL CONSTANTS OF METALLIC ABSORBERS	242
Metallic Media . . . Optical Constants . . . Normal Reflection . . . Reflection and Conductivity . . . Oblique Reflection . . . Comparison with Insulators . . . Total Reflection	
II. RADIATION AND SPECTRA	
CHAPTER 16. THERMAL RADIATION	259
Emissivity and Specific Intensity . . . Black-body Radiation . . . Selective Thermal Radiation . . . Color Temperature	

CONTENTS

ix

CHAPTER 17. LINE AND BAND SPECTRA	270
Discontinuous Spectra . . . Spectral Series . . . Wave Mechanics . . . Spectrochemical Analysis . . . Band Spectra . . . Rotational Bands . . . Vibration-rotation Bands . . . Electronic Bands . . . Microwave Absorption	
CHAPTER 18. X-RAY SPECTRA	299
Characteristic X-ray Emission and Absorption . . . X-ray Energy Levels . . . Non-diagram Lines . . . General Radiation	
CHAPTER 19. MAGNETO- AND ELECTRO-OPTICS	312
Faraday Effect . . . Kerr Magnetic Effect . . . Magnetic Double Re- fraction . . . Zeeman Effect . . . Stark Effect . . . Kerr Electro-optic Effect . . . Photoelectric Effects	
CHAPTER 20. MEASUREMENT OF COLOR	325
Trichromatic Colorimetry . . . Chromaticity . . . Color Mixture	
CHAPTER 21. THE SCATTERING OF LIGHT	331
Tyndall Effect . . . Rayleigh Scattering Formula . . . The Raman Effect	

IV. EXPERIMENTAL OPTICS

EXPERIMENT 1. CARDINAL POINTS OF A THICK LENS	338
2. ABERRATIONS OF LENSES	340
3. PHOTOMETRY AND ILLUMINOMETRY	344
4. COLORIMETRY	348
5. POLARIMETRY	349
6. SURFACE TESTING BY INTERFERENCE	353
7. DISPERSION AND THE PRISM SPECTRO- SCOPE	356
8. INDEX OF REFRACTION BY CRITICAL ANGLE	362
9. WAVE LENGTHS BY THE PLANE GRATING	365
10. MICHELSON INTERFEROMETER	368
11. FABRY-PEROT INTERFEROMETER	373
12. SPECTROPHOTOMETRY	377
13. INFRARED SPECTROMETER	380
14. CONCAVE GRATING	383
15. CHANNELLED SPECTRUM	388
16. EMISSION AND ABSORPTION IN THE ULTRAVIOLET	391
17. WAVE LENGTHS BY INTERPOLATION	393
18. QUANTITATIVE SPECTROCHEMICAL ANALYSIS	398
19. TRANSMISSIVITY BY THE SECTOR PHOTOMETER	401

EXPERIMENT 20. REFLECTION OF POLARIZED LIGHT	404
21. ELLIPTICALLY POLARIZED LIGHT	406
22. MEASUREMENT OF X-RAY WAVE LENGTHS	409
23. WAVE LENGTHS OF SOFT X-RAYS	415
24. CRYSTAL STRUCTURE BY THE POWDER METHOD	417
THE PHOTOGRAPHIC PROCESS	421
APPENDIX	431
Table 1. Fundamental Physical Constants	431
2. Some Useful Wave Lengths	432
3. Spectrum of the Mercury Arc in the Ultraviolet	433
4. Reflectivity	434
5. Indices of Refraction	434
6. Standard Visibility Function	436
ANSWERS TO PROBLEMS	437
BIBLIOGRAPHY	441
a. Geometrical Optics	441
b. Physical Optics	441
c. Radiation and Spectra	442
d. Experimental Optics	442
e. Laboratory Arts	442
INDEX	445

I. GEOMETRICAL OPTICS

Chapter 1. Fundamental Principles

1.1 Postulates and Basic Laws

Geometrical optics is primarily a theory of image formation by mirrors, prisms, and lenses, either singly or in various combinations. As such, it has an important place in all applications of optics, for example, in the use, selection, design, and manufacture of optical instruments. Its mathematical structure, which is largely geometrical or trigonometrical, is developed from a few fundamental principles, or postulates, that are generalizations drawn from experimental observation. They are (1) the rectilinear propagation of light, (2) the laws of reflection, (3) the laws of refraction, and (4) the assumption of the independence of light rays.

That light travels in straight lines in a homogeneous medium is assumed whenever one aims a gun, uses a surveying instrument, or makes predictions of eclipses. The accuracy obtainable in such cases is ample justification of the usefulness of this simplifying assumption. It is found convenient to describe rectilinear propagation of light by the use of light rays, which are lines drawn from a light source

in the direction of propagation of the light. These lines are straight if the medium is uniform. Each point of the source may be regarded as an origin of such rays, and it is common to differentiate among rays, ray pencils, and beams, as in Fig. 1.1(a), (b), and (c), respectively.

When a light ray meets an interface between two media, it generally separates into a reflected and a refracted ray as shown

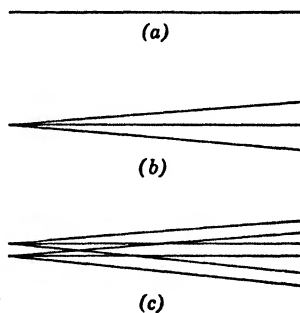


FIG. 1.1. The distinction between (a) a ray of light, (b) a pencil of rays, and (c) a beam of light.

in Fig. 1.2. Sometimes a high absorption coefficient makes it impossible to follow the refracted ray an appreciable distance into the refracting medium, as in the case of metals. If both media are isotropic, i.e., their physical properties are independent of direction, the laws of reflection and refraction are readily stated. The laws for anisotropic media will be discussed in the section dealing with crystal optics. The angles of the three rays in Fig. 1.2 are defined with respect to the normal to the interface. Thus a 0° angle of incidence means incidence along the normal. There are

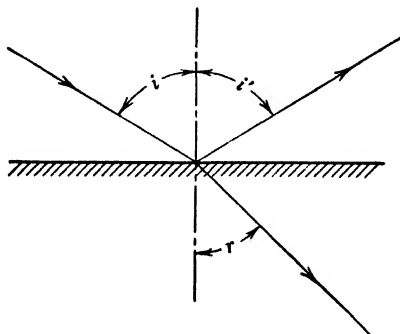


FIG. 1.2. Reflection and refraction at a plane interface between two media.

two laws of reflection: (1) the angle of reflection is equal to the angle of incidence, and (2) the reflected ray lies in the plane of incidence, this being the plane defined by the incident ray and the normal at the point of contact.

The laws of refraction for isotropic media are: (1) the ratio of the sine of the angle of incidence to the sine of the angle of refraction is constant, and (2) the refracted ray lies in the plane of incidence. The law of refracting angles is known as Snell's law

$$\frac{\sin i}{\sin r} = n$$

after Willebrord Snell, who first stated its equivalent in 1620. The constant n is called the **index of refraction** of the second medium with respect to the first. If the first medium is a vacuum, the value of n is called the **absolute index of refraction** of the refracting medium.

The fourth postulate, of the independence of light rays, overlooks all interference and diffraction effects. This is allowable in

most practical cases where the light waves possess a high degree of phase variability or "non-coherence" and where they are not restricted by small apertures. Although the fundamental principles of geometrical optics are drawn from experiment, one can easily set up conditions demonstrating the falseness of the first and the fourth postulates. The subject of physical optics abounds with illustrations of this fact. Nevertheless, circumstances are often such that one may safely use the principles of geometrical optics. They provide an effective, simplified method for solving many optical problems, and they form the basis of much of optical engineering. Geometrical optics may be considered to be an asymptotic form of wave optics with the wave length approaching zero. Newtonian mechanics bears an analogous relationship to wave mechanics. The usefulness of such asymptotic theories is well known; however, one must know the conditions under which they may be used without incurring large errors. This will be better understood after a study of physical optics. At this point it may suffice to state that it is safe to use the laws of geometrical optics to compute the location and size of any optical image of appreciable size with systems of usual apertures and, with additional information, to determine image brightness and field of view. It is also possible to compute defects of images, a matter of the utmost importance in the practical design of optical instruments.

1.2 Fermat's Principle

It is interesting to note that the four basic principles of geometrical optics are united in the principle of Fermat. Defining the optical length or optical path of a ray by the summation of the product of the index of refraction n and distance l along a light ray, Fermat's principle states that, when light passes from one point to another, it follows paths or "rays" which are lines connecting the two points in such a way that the optical path has a maximum or minimum value. In other words, the rays are defined by the condition that the variation in optical path is zero, i.e.,

$$\delta \sum n l = 0$$

for any infinitesimal deviation from the actual path. Rectilinear propagation in homogeneous media obviously follows from this principle, since a straight line is the shortest distance between

any two points. Since Fermat's principle is powerless to include any interference or diffraction effects, the principle of independence of light rays is automatically satisfied.

To demonstrate the derivation of the laws of reflection from Fermat's principle, consider S and Q and a reflecting plane P , Fig. 1.3. It is required to show that the shortest distance from S to Q by way of P satisfies the laws of reflection. In a plane through S and Q and perpendicular to the plane P , one constructs the normal N through S . On this normal a point S' is found which is

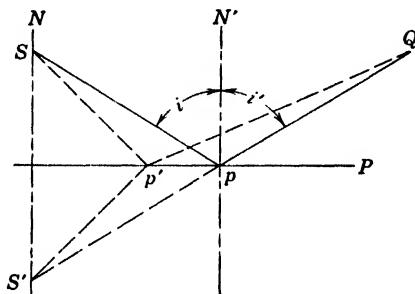


FIG. 1.3. Fermat's principle applied to reflection at a plane surface.

equally distant from P as the point S but lies on the other side as shown. The straight line $S'pQ$ is drawn, and then Sp . Since $Sp = S'p$ by the theorem referring to perpendicular bisectors, it follows that $SpQ = S'pQ$. Now, SpQ represents one possible path from S to Q by way of the reflecting surface. If any other path $Sp'Q$ is drawn, it is easy to show that this is equal to $S'p'Q$, and that this is greater than $S'pQ$, since the latter is a straight line joining the same two points. Therefore the light follows the path SpQ , which is the shortest. By erecting a normal N' at the point of reflection p and defining the angles of incidence i and of reflection i' as shown in Fig. 1.3, one can readily show, by comparing angles made at the parallel lines N and N' by the transversals Sp and $S'Q$, that $i' = i$. It is also quite evident that any path from S to Q by way of a point p'' on the plane P but outside the plane of the paper will be greater than the straight line $S'pQ$. Hence the reflected ray lies in the plane of incidence, this being the plane containing the incident ray Sp and the normal to the reflecting plane at P .

If the reflecting surface is any ellipsoid of revolution having S and Q as its foci, all paths SpQ will be equal to each other. Such

a surface is called an aplanatic surface. Any or all paths from S to Q by way of any point p on the surface may be taken by the light, which is then said to be focussed at Q .

If the reflecting surface has any general shape, it may be possible to find some aplanatic surface, a spheroid, with S and Q as foci, which is also tangent to the reflecting surface at some point p , as shown in Fig. 1.4. Now, if the given surface has a greater curvature at p than the aplanatic surface, the path from S to Q by way of any other point on it will be less than SpQ . In such a case,

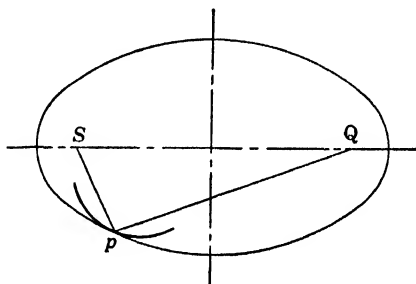


FIG. 1.4. The optical path SpQ is a maximum when reflection occurs at a surface having a greater curvature than an aplanatic surface which is tangent to it at p .

the only ray from S to Q by way of the reflecting surface is the one which takes the longest path SpQ . However, if the given surface is flatter than the aplanatic surface at the point of tangency p , the path followed will be the shortest. Hence one should think of Fermat's principle as one of extreme path, not necessarily of shortest path.

Analogous deductions can be made for refracting surfaces. Aplanatic refracting surfaces are discussed by Farwell in the *American Journal of Physics* for October 1941. They may be produced most practically by the molding of transparent plastics.

To derive the laws of refraction from Fermat's principle, consider the points S and Q , Fig. 1.5. on opposite sides of an interface between media whose indices of refraction are respectively n_1 and n_2 . To define a convenient coordinate system, first pass a plane through S and Q which is perpendicular to the interface. This plane is the plane of incidence, and its intersection with the interface is taken as the y axis. At some point along the y axis at which

a tentative light ray SOQ intersects it, erect the z axis normal to the interface, and take the x axis perpendicular to both y and z . The angles of incidence and refraction, i and r respectively, are defined as shown in Fig. 1.5. The relationship between these angles is determined by Fermat's requirement that

$$\delta \Sigma nl = 0$$

The path is varied by displacing the point at which the ray pierces the interface by amounts δx and δy . The resulting variation δL

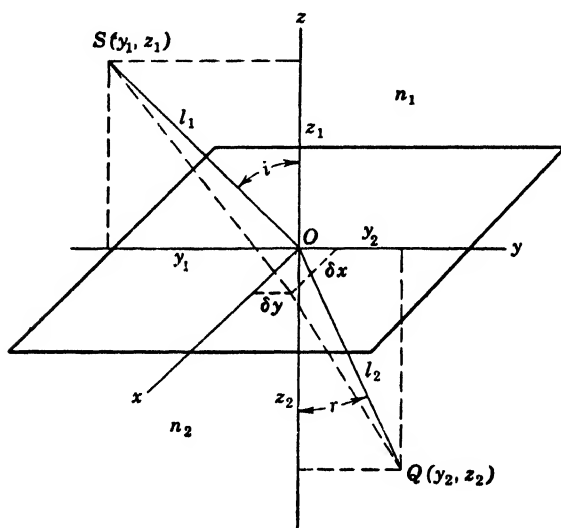


FIG. 1.5. Application of Fermat's principle to refraction at a plane interface.

in the optical path must be zero as far as infinitesimals of the first order are concerned. The variation in optical path is

$$\delta L = n_1(l_1 + \delta l_1) + n_2(l_2 + \delta l_2) - (n_1 l_1 + n_2 l_2)$$

By the usual expressions for distances in terms of Cartesian co-ordinates, this becomes

$$\begin{aligned} \delta L = & n_1 \sqrt{z_1^2 + (y_1 - \delta y)^2 + \delta x^2} + n_2 \sqrt{z_2^2 + (y_2 - \delta y)^2 + \delta x^2} \\ & - n_1 \sqrt{z_1^2 + y_1^2} - n_2 \sqrt{z_2^2 + y_2^2} \end{aligned}$$

Neglecting terms of the second order in δx and δy ,

$$\begin{aligned}\delta L &= n_1 \left[(z_1^2 + y_1^2) \left(1 - \frac{2y_1 \delta y}{z_1^2 + y_1^2} \right) \right]^{\frac{1}{2}} \\ &\quad + n_2 \left[(z_2^2 + y_2^2) \left(1 - \frac{2y_2 \delta y}{z_2^2 + y_2^2} \right) \right]^{\frac{1}{2}} \\ &\quad - n_1 (z_1^2 + y_1^2)^{\frac{1}{2}} - n_2 (z_2^2 + y_2^2)^{\frac{1}{2}} \\ &= - \left[\frac{n_1 y_1}{(z_1^2 + y_1^2)^{\frac{1}{2}}} + \frac{n_2 y_2}{(z_2^2 + y_2^2)^{\frac{1}{2}}} \right] \delta y \\ &= (n_1 \sin i - n_2 \sin r) \delta y\end{aligned}$$

since y_1 is negative. The variation is zero if

$$n_1 \sin i = n_2 \sin r$$

or

$$\frac{\sin i}{\sin r} = \frac{n_2}{n_1} = n$$

which is Snell's law. Moreover, since the variation in optical path depends only on second orders of the differentials δx , the extreme path will also lie in the plane of incidence. Thus all the laws of geometrical optics are contained in Fermat's principle.

It should be noted that, if the reflected or refracted ray is reversed in direction, the path followed after meeting the interface coincides with the original incident ray. This principle of reversibility is evident from Fermat's principle, which makes no distinction between directions of propagation. It is very useful in the solution of problems which may be more readily attacked from the reverse end.

1.3 Reflection and Refraction

When light meets a surface, it is divided into reflected and refracted rays. The relative amounts of energy in the two parts depend on the indices of refraction, the absorption coefficient, the angle of incidence, and the nature of polarization of the light. This topic will be discussed in detail later. At this point it will be noted that polished metals make excellent reflectors and that the small amount of refracted light is very rapidly absorbed. A metallic film must usually be very thin indeed to transmit any light.

Some metals, however, have wave-length bands of relatively high transparency. This is true of silver at about 3100 Å and of the alkali metals in the shorter ultraviolet. In these bands of transparency the reflectivity is low. Cesium has such a band near the visible region.

In optical applications of the laws of reflection, a metallic reflector may be used, or the surface of a non-metallic substance may be coated with a metallic film. The conditions for internal total reflection without the use of a metallic coating are discussed in Art. 15.7. Although aluminum does not reflect as well as polished silver, it is easily applied by evaporation *in vacuo* and is much more durable. Some substances, such as magnesium oxide or carbonate, are very efficient diffuse reflectors. This property is due to a combination of low absorption and a haphazard distribution of submicroscopic reflecting surfaces.

In optical applications of the laws of refraction, a transparent, homogeneous medium such as glass is used. If the angle of incidence is small, the fraction of the light, R , which is reflected at an interface between two media is given by Fresnel's formula, Art. 12.3,

$$R = \left(\frac{n - 1}{n + 1} \right)^2$$

where n is the index of refraction of the second medium relative to the first. For glass, whose index of refraction is about 1.5 relative to air, the reflection coefficient for light passing from air to glass is 0.04. Thus 96 per cent of the incident light is transmitted and refracted at each glass-air interface. The fraction of the light transmitted through a number N of such interfaces amounts to

$$T = (1 - R)^N$$

In a complicated instrument it is not unusual to find that T approaches 50 per cent. Methods for overcoming such large losses of light by the use of thin films on the surfaces will be discussed in Art. 10.6 after the basic principles have been established.

1.4 Absorption

When light of any homogeneous color passes through an absorbing medium, the intensity decreases according to Bouguer's law, more commonly known as Lambert's law:

$$I = I_0 e^{-ax}$$

In this formula x is the distance of propagation in the absorbing medium, a is the absorption coefficient, e is the base of natural logarithms, and I_0 is the intensity at $x = 0$. The transmissivity T for a thickness x is accordingly

$$T = \frac{I}{I_0} = e^{-ax}$$

For a unit thickness, $x = 1$, $T_1 = e^{-a}$, so that

$$T = T_1^x$$

This formula is convenient for computing the transmissivity for any thickness x when the unit transmissivity is known. Considering simultaneous reflection losses at N surfaces, each having a reflectivity R , and in a thickness x of absorber having a unit transmissivity T_1 , the combined transmissivity is

$$T = (1 - R)^N T_1^x$$

PROBLEMS

1. The glass window on a tank of water is 1 cm thick and has an index of refraction of 1.52. (a) If the angle of incidence in air is 60° , what is the angle of refraction in the water? (b) What is the angle of deviation?

2. Let the x, y plane coincide with the interface between air and water. If a light ray passes from a point whose Cartesian coordinates x, y, z are 6, 1, 5 to a point whose coordinates are $-8, 2, -6$, where will it pierce the surface?

3. A gun is to be aimed at an object submerged in water 30 ft away horizontally, and apparently (image) on a line making an angle of 10° downward from the horizontal. (a) If the muzzle of the gun is 4 ft above the surface of the water, at what angle must it be aimed to point directly at the object? (b) How far is the object below the surface of the water?

4. What fraction of the light is transmitted by an optical instrument containing twelve glass-air interfaces ($n = 1.5$) and a total path in glass of 10 cm, the absorption in the body of the glass being 1 per cent per centimeter?

Chapter 2. Ideal Optical Systems

2.1 Images

When light passes from a point P to a point Q by way of one or more reflecting or refracting surfaces, Fermat's principle in general gives a unique direction for the ray arriving at Q . Sometimes, however, a pencil of rays may focus at some particular point Q , as in the case of any portion of a surface which is aplanatic with respect to P and Q . If the rays actually meet at Q so as to produce there a concentration of light which may be received on a screen, the point Q is said to be a **real focus** or a **real image** of the point P . If the point Q is merely the intersection of a backward extension of diverging light rays, it is termed a **virtual focus** or a **virtual image** of P .

Because of the principle of reversibility, the light rays may be traced in the reverse direction with the result that P is the image of Q . Hence P and Q are said to be **conjugate points** or conjugate foci.

The reader is familiar with the application of these ideas to the mirror and to the thin lens. In either case, one may express the relation between the conjugate distances s and s' measured from the mirror or the thin lens by the equation

$$\frac{1}{s} + \frac{1}{s'} = \frac{1}{f}$$

where f is the **focal length** or distance from the mirror or thin lens to the point at which the image of a very distant object is found, i.e.,

$$f = \lim_{s \rightarrow \infty} s'$$

For a spherical mirror $f = R/2$, while for a thin lens

$$\frac{1}{f} = (n - 1) \left(\frac{1}{R_1} - \frac{1}{R_2} \right)$$

where n is the relative index of refraction of the material of the

lens with respect to the surrounding medium, and the R 's are the radii of curvature of its two faces. The sign of R for a mirror is positive if the surface is concave, while the signs of R_1 and R_2 are positive for surfaces which are convex toward the incident light. The conjugate distances s and s' are positive for a real focus and negative for a virtual focus. The size of the image is ordinarily found by multiplying the size of the object at P by the transverse magnifying power β , where

$$\beta = \left| \frac{s'}{s} \right|$$

In order to associate a positive magnification with an erect image and a negative magnification with an inverted image, one writes

$$\beta = - \frac{s'}{s}$$

Other sign rules are used by some writers, but the formulas given above require the sign rules just stated. If a contrary sign rule is adopted for any quantity, its sign in each formula is reversed.

2.2 Collinear Transformation

The above equations are derived in every elementary textbook on light by the use of the laws of reflection and refraction, with the assumption that all angles are small enough so that one may write $\sin \theta = \theta$ and $\cos \theta = 1$. Abbe was the first to show that these equations may be derived without the use of any physical laws such as the laws of reflection and refraction, but in a very general manner by the use of pure algebra. This fact makes it not at all surprising that images formed by any combination of reflecting and refracting surfaces, and even by electron lens systems, can be computed by the formulas to be derived. Thus Abbe's theory provides the best analytical approach to the theory of all kinds of optical systems, particularly actual thick lenses. It indicates the general properties of such systems and defines the significant constants.

For Abbe's method of analysis of an optical problem, one adopts two systems of Cartesian coordinates: x, y, z for object points or points in an "object space," and x', y', z' for the corresponding points in an "image space." These spaces overlap everywhere, since objects and images may be virtual as well as real. An **ideal**

optical system is defined as one which will produce respectively one point, straight line, or plane in the image space for any one point, straight line, or plane in the object space. In such a system the relations between coordinates in the two spaces can be expressed mathematically by the collinear transformation:

$$\begin{aligned}x' &= \frac{a_1x + b_1y + c_1z + d_1}{ax + by + cz + d} \\y' &= \frac{a_2x + b_2y + c_2z + d_2}{ax + by + cz + d} \\z' &= \frac{a_3x + b_3y + c_3z + d_3}{ax + by + cz + d}\end{aligned}$$

It is not difficult to prove that this transformation is of the point-to-point, line-to-line, and plane-to-plane type by applying it successively to the equations of a point, plane, or line (two intersecting planes).

An optical system is usually simpler than the above equations indicate, since it generally has an axis of symmetry. If this is so, the axis of symmetry may be taken as the x and x' axis, and any plane through this axis may be taken as the x, y plane. In this plane $z = z' = 0$, and the above equations reduce to

$$\begin{aligned}x' &= \frac{a_1x + b_1y + d_1}{ax + by + d} \\y' &= \frac{a_2x + b_2y + d_2}{ax + by + d}\end{aligned}$$

Furthermore, the value of x' must be independent of y since the location of an image in an ideal system does not depend on the size of the object; and y' must simply reverse its sign with no change in magnitude when y reverses in sign. These requirements make $b_1 = b = 0$ and $a_2 = d_2 = 0$, respectively. The resulting transformation then has the simple form:

$$\begin{aligned}x' &= \frac{a_1x + d_1}{ax + d} \\y' &= \frac{b_2y}{ax + d}\end{aligned}$$

Solving these for x and y in terms of x' and y' , one finds that

$$x = \frac{d_1 - dx'}{ax' - a_1}$$

$$y = \frac{y'(ad_1 - a_1d)}{b_2(ax' - a_1)}$$

These equations give the relation between the coordinates of conjugate points in an ideal optical system which has an axis of symmetry. The values of the constants depend on the nature of the system as well as on the choice of the origins of the two systems of coordinates. Except for the condition that the origins be located on the axis of the optical system, their position has so far been unspecified. On investigating the properties of the ideal system, certain cardinal points will be discovered which make more suitable origins for the systems of coordinates.

2.3 Principal Foci and Principal Planes

It is evident that, when $x = -d/a$, the values of x' and y' become infinite. The equation $ax + d = 0$ then defines a plane in the object space which is perpendicular to the x axis, each point in this plane having its conjugate at infinity. This is the **principal focal plane** in the object space, and its intersection with the x axis defines the **principal focus** in this space. This point is also called the **first principal focus**. Similarly $ax' - a_1 = 0$ defines the principal focal plane in the image space, and its intersection with the x' axis at F' is called the **second principal focus**. Note that the positive directions of the x and x' axes are conventionally taken in opposite directions, as shown in Fig. 2.1.

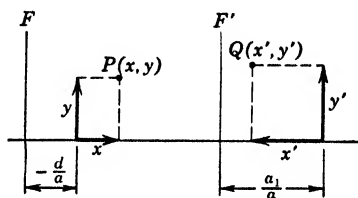


FIG. 2.1. Cartesian coordinates of object space and image space, showing location of principal focal planes at F and F' .

The general equations reduce to a simpler form if these principal foci are taken as the respective origins in the object space and in

the image space. The transformation is accomplished by making the substitutions

$$x = x_0 - \frac{d}{a}; \quad y = y_0$$

and

$$x' = x_0' + \frac{a_1}{a}; \quad y' = y_0'$$

After algebraic reduction, one finds that

$$x_0 x_0' = \frac{d_1 a - a_1 d}{a^2}$$

and

$$\frac{y'}{y} = \frac{y_0'}{y_0} = \frac{b_2}{a x_0}$$

These equations can be simplified in form by introducing two new constants f and f' defined by the equations

$$\frac{b_2}{a} = f \quad \text{and} \quad \frac{d_1 a - a_1 d}{a^2} = f f'$$

Since the former origin of coordinates will not be used again, the subscripts of x_0 and x_0' can be dropped without ambiguity, so that

$$x x' = f f'$$

and

$$\beta \equiv \frac{y'}{y} = \frac{f}{x} = \frac{x'}{f'}$$

In these equations x and x' refer to distances measured from the principal foci in the object and image space, respectively. The equations suffice to determine the location, x' , and the transverse size, y' , of the image of any object once the constants f and f' are known. These may be determined in the laboratory by fitting experimental data to the above equations, or they may be computed by formulas derived by the use of the physical laws of reflection and refraction or their equivalent, as shown in Chapter 3.

The significance of f and f' can best be appreciated by defining the **unit planes** of an optical system as those conjugate planes for which the transverse linear magnification $y'/y = +1$. It is readily seen that this yields $x = f$ as the defining equation of the **first**

principal plane and $x' = f'$ as the equation of the **second principal plane**. Their intersections with the axis of symmetry are the **principal points**. The distances between the principal points and the associated principal foci are the focal lengths of the optical system. These relations are shown in Fig. 2.2.

The conjugate object and image locations are often given by the respective distances s and s' from the principal planes. According to the usual rule of signs, s and s' are measured as positive in opposite directions to x and x' as shown in Fig. 2.3, where all quan-

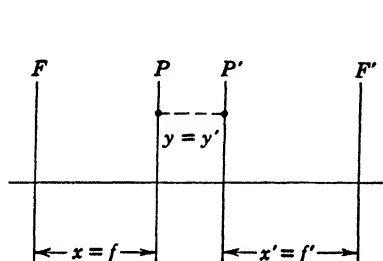


FIG. 2.2. Location of principal planes P and P' , showing their property of unit transverse magnification.

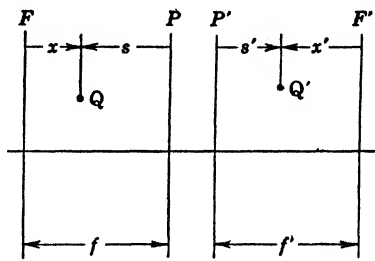


FIG. 2.3. Principal planes used as origins for measurement of the conjugate distances s and s' .

tities are positive. It is evident from this figure that $x = f - s$ and $x' = f' - s'$, so that

$$xx' = ff' - sf' - fs' + ss' = ff'$$

from which it follows that

$$\frac{f}{s} + \frac{f'}{s'} = 1$$

In addition, it can be shown that the transverse magnification

$$\frac{y'}{y} = -\frac{s'f}{sf'}$$

If $f = f'$, these equations reduce to the familiar lens equations

$$\frac{1}{s} + \frac{1}{s'} = \frac{1}{f}$$

and

$$\frac{y'}{y} = -\frac{s'}{s}$$

When the equations for f and f' are derived from physical laws, it will be proved that they are in the ratio of the indices of refraction in the object and image spaces; thus

$$\frac{f'}{f} = \frac{n'}{n}$$

Therefore any optical system in air, or with identical media on its entrance and exit sides, will obey the elementary lens equations, with the requirement, however, that the distances s and s' be measured from the respective principal planes. The locations of

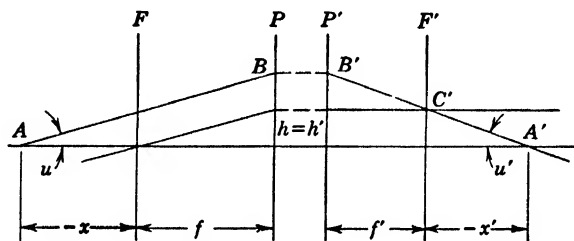


FIG. 2.4. Tracing of the conjugate of ray AB by the use of the properties of the principal planes and the focal plane F' .

these planes must also be derived by the use of physical laws or by experiment.

The properties of the principal planes and focal planes may be used to trace rays through an optical system. Thus in Fig. 2.4, let AB represent any ray in the object space. Its conjugate is fixed by locating the conjugates of any two points on it. One of these is B' , the conjugate of B , which must be at the same height $h' = h$ above the axis in the second principal plane. Another point is found by tracing an auxiliary ray parallel to AB and passing through the first principal focus. This ray will emerge parallel to the axis in the image space. It intersects the F' plane at the point C' . This point must also lie on the conjugate ray to AB , since parallel rays in the object space must intersect in the focal plane F' . Thus $B'C'$ defines the emergent ray which is conjugate to AB .

2.4 Angular Magnification and the Nodal Points

The angle that a ray makes with the principal axis is called its **slope angle**. The ratio of the tangents of conjugate slope angles is called the **angular magnification**. The signs of the slope angles

u and u' are considered positive when the rays slope as shown in Fig. 2.4, in which x and x' are negative. Thus, referring to the figure, the angular magnification is found to be

$$\gamma \equiv \frac{\tan u'}{\tan u} = \frac{h'(f-x)}{(f'-x')h} = -\frac{f}{x'} = -\frac{x}{f'}$$

The **nodal points** of an optical system are defined as those conjugate points on the axis through which conjugate rays are parallel to each other. In view of our rule of signs, this means that the angular magnification $\gamma = -1$. Hence the locations of the nodal points relative to the principal foci are given by the equations

$$x = f' \quad \text{and} \quad x' = f$$

When the indices of refraction of the media in the object and image spaces are the same, $f' = f$, and the nodal points coincide with the principal points. In any case, the distance between the nodal points is equal to the distance between the principal points.

The two principal foci, the two principal points, and the two nodal points are called the **cardinal points** or the **Gauss points** of an optical system. They are extremely useful in discussing the properties of optical systems and in making computations, but it should not be overlooked that they were derived for an ideal optical system, and that actual systems will deviate more or less from the exact form of this theory. However, the first order theory of actual systems, in which one deals with small angles and may freely approximate $\sin \theta = \theta$ and $\cos \theta = 1$, agrees in form with the theory of the ideal optical system just given.

2.5 Classification of Optical Systems

Optical systems are classified according to the signs of f and f' , these being determined by the sequence of the cardinal points $FP'P'F'$ as one proceeds along the principal axis in the general direction of the incident light. If P follows F , the value of f is positive and the system is said to be **convergent**, regardless of the sign of f' . If F' follows P' , then f' is also positive and the system is **dioptric convergent**, whereas if P' follows F' , then f' is negative and the system is **katoptric convergent**. If f is negative, the system is **divergent**; if f' is also negative, it is **dioptric divergent**, whereas if f' is positive, it is **katoptric**.

Thus, like signs of f and f' refer to a dioptric system and unlike signs refer to a katoptric system. The former is also said to be a **concurrent** system since any longitudinal displacement dx of the object results in a displacement dx' of the image which is in the same direction along the axis. This is readily seen by deriving an expression for the longitudinal magnification α

$$\alpha \equiv \frac{dx'}{dx} = -\frac{ff'}{x^2} = -\frac{x'^2}{ff'}$$

Because of the squared term in x or x' , this magnification along the axis is always negative if f and f' have the same sign. Because positive x and x' are in opposite directions, this means that the changes in x and x' are concurrent in space when α is negative. Similarly, one sees that focal lengths of opposite sign refer to a **contracurrent** system.

The three magnifications α , β , and γ are obviously not independent. The following relations between them are readily derived from the equations for these quantities:

$$\alpha = \frac{\beta}{\gamma} \quad \text{and} \quad \beta\gamma = -\frac{f}{f'}$$

An important practical case of the optical system, the telescopic system, has its principal planes and focal planes at infinity. Moreover, the focal lengths are also infinite (Art. 3.7). Many of the equations derived for the ideal optical system will give indeterminate results when applied to this special case. These results may, however, often be evaluated by starting with a large f and f' and passing to the limit. Thus, for example, the last equation above reduces to

$$\beta\gamma = -1$$

when the object space and the image space are filled with a medium of the same index of refraction, which is often the case. Other properties of the telescopic system will be discussed in Art. 3.7.

PROBLEMS

1. (a) Prove that a magnification M is obtained when the image is projected on a screen at a distance of $M + 1$ focal lengths from the second principal plane of the lens. (b) What is the corresponding object distance?

By Snell's law, Art. 1.2,

$$\sin \phi' = \frac{n}{n'} \sin \phi \quad (2)$$

From the computed values of ϕ and ϕ' , one obtains the slope angle

$$u' = \phi - u - \phi' \quad (3)$$

The conjugate distance s' is then found by applying the sine law to the triangle RCQ'

$$\frac{s' - r}{\sin \phi'} = \frac{r}{\sin u'} \quad (4)$$

giving for s'

$$s' = r + \frac{r \sin \phi'}{\sin u'} \quad (4')$$

This sequence of formulas enables one to fix a ray after refraction at the first of a series of spherical surfaces when the incident ray is fixed by its s and u . The same procedure may then be applied at the second surface, and so on, through the entire optical system. A neatly tabulated system of computation will facilitate increased speed and accuracy. The signs of all the quantities must be carefully observed, as failure to notice these signs is a frequent source of error. For certain special cases, such as that of a plane refracting surface, or in tracing a ray parallel to the axis, the formulas for u' and s' become indeterminate. They must accordingly be replaced by other formulas which, however, are easily derived, and are left to the student as an exercise. Note that a ray parallel to the axis is completely defined by just its distance h above the axis instead of the usual two quantities s and u .

Trigonometrical ray tracing is an extremely important and common operation in the design of optical systems; hence the student should obtain some practice with it. For further discussion of this subject, consult Conrady's *Applied Optics*, or see Hardy and Perrin's *The Principles of Optics*. Note that equations (1) and (4) may be combined to give

$$\frac{s + r}{s' - r} = \frac{n' \sin u'}{n \sin u} \quad (5)$$

which may be used as a computational check.

3.2 First Order Theory

So far, no approximations have been made, and the results of the computations would, in general, indicate considerable deviation from an ideal optical system. However, the ideal system will be closely approximated if the angles u and u' and ϕ and ϕ' are small. In that event, the rays are said to be **paraxial**, and one finds that

$$\frac{\sin u'}{\sin u} \doteq \frac{PR}{s'} \frac{s}{PR} = \frac{s}{s'}$$

and (5) becomes

$$\frac{s + r}{s' - r} = \frac{n' s}{n s'}$$

This is easily converted into the form

$$\frac{n}{s} + \frac{n'}{s'} = \frac{n' - n}{r} \quad (6)$$

or

$$\frac{f}{s} + \frac{f'}{s'} = 1 \quad (7)$$

where $f = nr/(n' - n)$ and $f' = n'r/(n' - n)$.

Thus in (7) we obtain the same formula as for an ideal optical system in which the constants f and f' may be identified as the two focal lengths, provided that s and s' may be shown to be measured from the principal planes. One may locate P' by the condition that it must always pass through a point on the refracted ray which is at the same height h' as the height of its conjugate point h on the incident ray. Taking an incident ray which is parallel to the axis so that h is the same everywhere, one finds that the condition $h = h'$ shows that P' must be on the spherical refracting surface. For the paraxial region of this surface P' approaches the vertex V , so that the principal plane P' is tangent to the surface at V . By reversing this procedure one may show that P is also at V . Moreover, (6) indicates that, when $s = \infty$, $s' = n'r/(n' - n)$ locates the principal focus F' at that distance to the right of the vertex V . Similarly, F is located at the distance $nr/(n' - n)$ to the left of the vertex. These distances are accordingly the focal lengths f' and f respectively. The nodal points N' and N may be shown to coincide with each other at the center of curvature C

either by the use of the condition $\phi = \phi'$ or by the use of $x = f'$ and $x' = f$.

The formula for reflection of paraxial rays at a spherical mirror may be obtained from the above by using the artifice that $n = -n'$. This makes $f = -r/2$ and $f' = +r/2$ for a convex surface, the conventions for the signs of s , s' , and r being unchanged. It is customary, however, to reverse the sign convention for s' and r when the formula is applied to mirrors, so that the formula becomes

$$\frac{1}{s} + \frac{1}{s'} = \frac{2}{r}$$

To obtain the formula for the transverse magnification by a single refracting surface, one may substitute the values for f and f' in the general formula for an ideal optical system. Thus

$$\beta = \frac{y'}{y} = -\frac{s' f}{s f'} = -\frac{s' n}{s n'}$$

For reflectors, this becomes

$$\beta = -\frac{s'}{s}$$

using the altered sign convention for s' .

Again, in the case of a single refracting surface, Fig. 3.1, it is seen that the angular magnification is

$$\gamma = \frac{\tan u'}{\tan u} = \frac{s}{s'}$$

Combining this with the above formula for β , one finds, using the original sign convention, that

$$\beta = \frac{y'}{y} = -\frac{\tan u}{\tan u'} \frac{n}{n'}$$

or

$$ny \tan u = -n'y' \tan u'$$

which is Lagrange's law. The negative sign disappears if the slope angles are defined as positive when the rays travel at acute angles measured counterclockwise from the principal axis in the direction of the light. The quantity $ny \tan u$ is a useful optical invariant which reduces to nyu when the slope angle is small.

It is evident that if $ny \tan u$ is invariant for refraction at one spherical surface, it will be invariant at the next, and so on, through any number of coaxial spherical surfaces. This leads to a useful general relation between the linear and angular magnification, namely,

$$\frac{n'y' \tan u'}{ny \tan u} = \frac{n'}{n} \beta\gamma = -1$$

or

$$\beta\gamma = -\frac{n}{n'}$$

using the original sign conventions. Since it was proved that for an ideal optical system

$$\beta\gamma = -\frac{f}{f'}$$

it follows that the focal lengths are directly proportional to the indices of refraction on the two sides of any optical system. This is, of course, evident for a single refracting surface from the formulas for f and f' , but this discussion using Lagrange's law shows that it also applies to any number of surfaces, hence to any optical system.

3.3 Combination of Optical Systems

A **simple lens** consists of two spherical refracting surfaces. Its cardinal points may be derived from those of the two surfaces. As a foundation for this derivation, a general combination of any two optical systems will be first considered. The problem is to locate the cardinal points F , P , P' , F' of the combination when those of each system F_1 , P_1 , P_1' , F_1' and F_2 , P_2 , P_2' , F_2' are given. Let Δ represent the separation $F_1'F_2$, which is considered to be positive if the order of these points with respect to the general direction of the light is as shown in Fig. 3.2. To locate the principal focus of the combination, one needs only to trace any incoming ray parallel to the axis as shown, and to locate the intersection F' of its conjugate after passing through both systems. The second principal plane P' is found by extending this ray and the original incident ray until they cross at a height $h' = h$ above the axis. The axial distance $F'P'$ is the focal length f' of the system and is negative in this example. A reverse of this pro-

cedure locates F and P and defines the first focal length f . From this graphical solution of the problem one may derive formulas for the computation of f and f' and for the location of the principal planes. The latter can be measured from any convenient point or points, but usually are fixed by the distances p and p' defined in

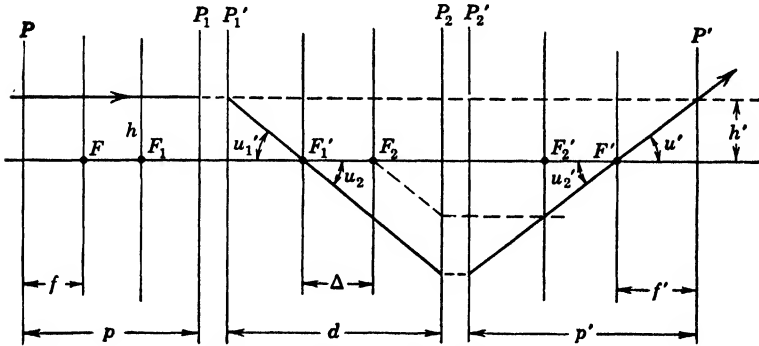


FIG. 3.2. Combination of two optical systems.

Fig. 3.2. These distances are to be considered positive when the terminal points are in the order shown in the figure.

To derive the necessary formulas, the points F and F' will first be located by the use of the conjugate focal relation

$$x_2' = \frac{f_2 f_2'}{x_2}$$

so that

$$F_2' F' = \frac{f_2 f_2'}{\Delta}$$

Similarly

$$F_1 F = \frac{f_1 f_1'}{\Delta}$$

In the substitution for x_2 and x_2' it should be observed that both quantities are negative and therefore the algebraic sign cancels. The distances of the principal points P and P' from the principal foci F and F' , respectively, are determined by solving for the focal lengths f and f' . Thus

$$f' = \frac{h'}{\tan u'}$$

where f' and u' are both negative in Fig. 3.2. Since

$$\tan u' = \tan u_2'$$

and

$$\tan u_2' = \gamma_2 \tan u_2 = -\gamma_2 \tan u_1'$$

it follows that

$$\tan u' = \frac{x_2}{f_2'} \frac{h'}{f_1'} = - \frac{h' \Delta}{f_1' f_2'}$$

Hence

$$f' = - \frac{f_1' f_2'}{\Delta}$$

and similarly

$$f = - \frac{f_1 f_2}{\Delta}$$

Formulas in which the separation of the two systems is given by $d = P_1' P_2$ are generally more convenient to use. Referring to Fig. 3.2, it is seen that

Thus

$$\Delta = d - f_1' - f_2$$

$$f' = - \frac{f_1' f_2'}{d - f_1' - f_2} \quad (8)$$

and

$$f = - \frac{f_1 f_2}{d - f_1' - f_2} \quad (9)$$

Note that, if the medium on both sides of the lenses is the same, formulas (8) and (9) may be written in the common form

$$\frac{1}{f} = \frac{1}{f_1} + \frac{1}{f_2} - \frac{d}{f_1 f_2}$$

This formula by itself has a very restricted usefulness unless the principal planes are known. Referring again to Fig. 3.2, it is seen that

$$p' = f_2' + F_2 F' - f'$$

so that

$$p' = f_2' + \frac{f_2 f_2' + f_1' f_2'}{\Delta}$$

and

$$p' = \frac{f_2' d}{d - f_1' - f_2} \quad (10)$$

Similarly, it can be shown that

$$p = \frac{f_1 d}{d - f_1' - f_2} \quad (11)$$

The set of formulas (8), (9), (10), and (11), which apply to the combination of any two optical systems, will now be applied to the problem of a simple lens in air.

3.4 The Simple Lens in Air

Consider two spherical surfaces with an intervening medium of index n , and take the line through their centers of curvature as the

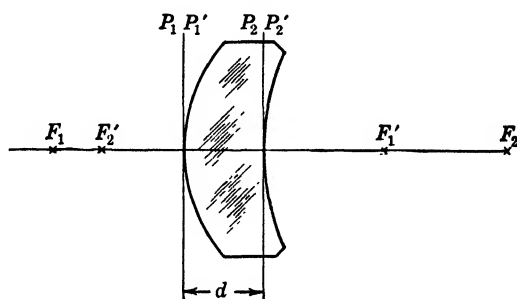


FIG. 3.3. Simple thick lens as a combination of two refracting spherical surfaces.

principal axis of the simple lens obtained in this way. The radii of curvature of the two surfaces, r_1 and r_2 , are both positive when convex toward the light, as in Fig. 3.3. The principal planes and foci of each single surface and of the combination are indicated by the usual symbols. The focal lengths of the components are

$$f_1 = \frac{r_1}{n - 1}; \quad f_1' = \frac{nr_1}{n - 1}$$

$$f_2 = -\frac{nr_2}{n - 1}; \quad f_2' = -\frac{r_2}{n - 1}$$

All the general equations contain the same denominator:

$$d - f_1' - f_2 = d - \frac{n(r_1 - r_2)}{n - 1} = \frac{(n - 1)d - n(r_1 - r_2)}{n - 1}$$

Substitution into the general equations gives the results:

$$f' = f = - \frac{nr_1r_2}{(n-1)[n(r_1-r_2) - (n-1)d]}$$

$$p = - \frac{r_1d}{n(r_1-r_2) - (n-1)d}$$

$$p' = + \frac{r_2d}{n(r_1-r_2) - (n-1)d}$$

One may write the expression for the power of a simple lens in the form

$$\frac{1}{f'} = \frac{1}{f} = (n-1) \left[\frac{1}{r_1} - \frac{1}{r_2} + \frac{(n-1)d}{nr_1r_2} \right]$$

For a lens which is thin enough so that $(n-1)d \ll n(r_1-r_2)$, an examination of the foregoing equations shows that they may be approximated by

$$\frac{1}{f} = (n-1) \left(\frac{1}{r_1} - \frac{1}{r_2} \right)$$

$$p = \frac{r_1d}{n(r_2-r_1)}$$

$$p' = - \frac{r_2d}{n(r_2-r_1)}$$

For example, a thin equiconvex lens, whose $n = 1.5$ and $r_2 = -r_1$, has its principal planes located at approximately $p = -d/3$ and $p' = -d/3$, as shown in Fig. 3.4(a). A plano-convex lens with $r_2 = \infty$ and $n = 1.5$ has its principal planes at $p = 0$ and $p' = -2d/3$, as shown in Fig. 3.4(b). It is instructive to examine other cases. The fact that the formulas hold only for rather thin lenses should be kept in mind. The location of the principal planes for paraxial rays passing through a sphere, using the exact formulas, will emphasize the magnitude of the approximation resulting from the dropping of $(n-1)d$ when the lens is not thin.

The following convenient formulas for the location of the principal planes may be readily obtained from the preceding exact forms without approximation:

$$p = - \frac{f}{f_2} d; \quad p' = - \frac{f'}{f_1'} d$$

These may be applied to a combination of optical systems as well as to a simple thick lens.

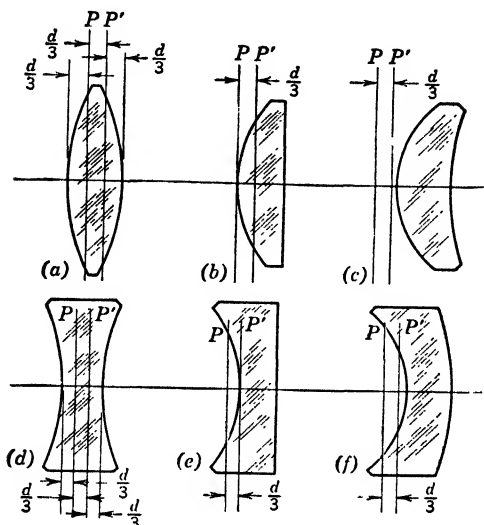


FIG. 3.4. Relative locations of principal planes in six common lenses of small thickness: (a) equiconvex, (b) plano-convex, (c) convex meniscus, (d) equiconcave, (e) plano-concave, (f) concave meniscus.

3.5 Dioptric Power of a Lens

Instead of focal lengths, one may use the powers of the lenses or lens surfaces in diopters, this being the reciprocal of the focal length in meters multiplied by the index of refraction of the medium in which that focal length is measured. For the surfaces of a thick lens, one has the powers

$$D_1 = \frac{1}{f_1} = \frac{n}{f_1'} \quad \text{and} \quad D_2 = \frac{1}{f_2'} = \frac{n}{f_2}$$

The power of the lens is then

$$D = D_1 + D_2 - D_1 D_2 \frac{d}{n}$$

and the principal planes are at

$$p = -\frac{d}{n} \frac{D_2}{D} \quad \text{and} \quad p' = -\frac{d}{n} \frac{D_1}{D}$$

In dealing with ophthalmic lenses, it is important to know the vertex focal length and the vertex power. The **vertex focal length** f_v is the distance of the principal focus measured from the rear vertex of the lens. Hence

$$f_v = f' + p' = f' \left(1 - \frac{d}{f_1'} \right)$$

The **vertex power** D_v is the reciprocal of the vertex focal length. Therefore

$$D_v = \frac{1}{f'} \left(\frac{1}{1 - \frac{d}{f_1'}} \right) = D \left(\frac{1}{1 - \frac{d}{f_1'}} \right) = D \left(\frac{1}{1 - \frac{d}{n} D_1} \right)$$

where D is the ordinary or optical power. When linear distances in these formulas are in meters, D and D_v are in units called **diopters**. It is evident that D_v differs appreciably from D whenever f_1' is not many times greater than d ; otherwise the difference is slight.

The powers of the surfaces D_1 and D_2 may be measured by a Geneva lens gauge or other form of spherometer. In the Geneva gauge there are three projecting pins which make contact with the lens surface. The outer two are fixed, while the inner one is connected to a pointer and scale by means of a mechanical lever system. The scale usually reads diopters $\left(D = \frac{n-1}{r} \right)$ corresponding to an index of refraction of 1.530. If the glass has a different index of refraction, one must multiply the gauge reading by the ratio of $(n-1)$ to 0.530. If the lens thickness is negligible, the lens power is simply the algebraic sum of the corrected readings for the two faces.

3.6 Alternative Computation of Focal Length

The focal length of an optical system consisting of any number of spherical surfaces with their centers of curvature on a common axis may be readily computed from the conjugate distances s and s' for the separate surfaces. These are obtained by a step-by-step application of the conjugate distance equation to the consecutive surfaces. For the first of the surfaces, the transverse magnification

$$m_1 = - \frac{n s_1'}{n_1 s_1}$$

and similarly for the others. The overall magnification

$$\begin{aligned} m &= m_1 m_2 m_3 \dots = \left(-\frac{n}{n_1} \frac{s_1'}{s_1} \right) \left(-\frac{n_1}{n_2} \frac{s_2'}{s_2} \right) \left(-\frac{n_2}{n_3} \frac{s_3'}{s_3} \right) \dots \\ &= -\frac{n}{n'} \frac{s'}{s} \end{aligned}$$

As s_1 , and therefore s , approaches infinity, s' approaches f' , and s_1' approaches f_1' . Then

$$f' = f_1' \left(-\frac{s_2'}{s_2} \right) \left(-\frac{s_3'}{s_3} \right) \dots$$

3.7 Telescopic System

By a telescope, one usually means any combination of lenses used for viewing an object which is at least several focal lengths from the first lens. However, in this section, a **telescopic system**

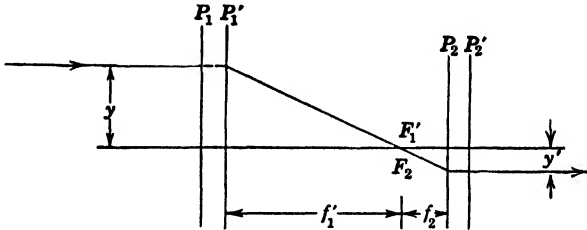


FIG. 3.5. Telescopic system.

will be defined as a combination of two systems with a separation $\Delta = 0$. Such a telescopic system is illustrated in Fig. 3.5. The subscript 1 refers to the objective system, and 2 to the eyepiece. Both of these are usually compound lenses for reasons to be explained in the next chapter.

Since $\Delta = 0$ in the telescopic case, the values of f , f' , p , and p' obtained from the formulas in Art. 3.3 are all infinite, and the cardinal points F , F' , P , P' are at infinity. Nevertheless, it will be shown that the longitudinal, lateral, and angular magnifications are all finite. It is on account of this property that the telescopic system is so useful and deserves further discussion.

Referring to Fig. 3.5, it is seen that the transverse magnification is, by similar triangles,

$$\beta = \frac{y'}{y} = -\frac{f_2}{f_1'}$$

The parallel incident and emergent rays traced in Fig. 3.5 define the height of any object and its image, since the conjugate points must be somewhere on the conjugate rays. Hence the transverse magnification is independent of the conjugate distances.

Since, in general,

$$\beta\gamma = -\frac{f}{f'} = -\frac{n}{n'}$$

it follows that the angular magnification

$$\gamma = -\frac{1}{\beta} \frac{n}{n'} = \frac{f_1' n}{f_2 n'}$$

This is also constant, regardless of conjugate distances. Moreover, since the longitudinal magnification is

$$\alpha = -\frac{ff'}{x^2} = -\beta^2 \frac{n'}{n}$$

this is also constant. The indices of refraction may be eliminated from these equations by making use of

$$\frac{f}{f'} = \frac{f_1 f_2}{f_1' f_2'} = \frac{n}{n'}$$

Thus one obtains

$$\alpha = -\frac{f_2 f_2'}{f_1 f_1'}$$

and

$$\gamma = \frac{f_1}{f_2'}$$

If there is air on both sides of the telescopic system, the formulas reduce to

$$\alpha = -\left(\frac{f_2}{f_1}\right)^2; \quad \beta = -\frac{f_2}{f_1}; \quad \gamma = +\frac{f_1}{f_2}$$

These are not only finite, but are independent of the object distance x or s . The magnifying power γ is the commonly designated power of the telescope. It is the ratio of the tangents of the visual angles which are subtended by the image seen from the position of the eye, and by the object as seen from the center of the objective, the latter angle being nearly equal to that subtended at the unaided eye.

PROBLEMS

1. A biconvex lens having radii of curvature of 150 mm and -75 mm, a thickness of 5 mm, and a diameter of 30 mm is made of crown glass having the indices of refraction given in Table 5(*d*) in the Appendix. Determine by trigonometrical ray tracing the foci and slope angles of a rim ray, a $\frac{2}{3}$ rim ray, and a paraxial ray for the principal Fraunhofer lines, assuming that the incident rays are parallel to the axis of the lens. (*Since a great deal of computation is required for a complete solution, it is advisable to assign parts of this problem to various members of the class.*)

2. A real image is formed by a concave mirror 60 cm away from the image. If the image is five times as large as the object, (*a*) how far is the object from the mirror, and (*b*) what is the radius of curvature of the mirror?

3. In order to find the radius of curvature of a convex mirror, a small light source is placed 10 cm in front of it, and the light is reflected into a concave mirror of 10-cm focal length which forms a real image 20 cm away from it. If the two mirrors are 15 cm apart, what is the radius of curvature of the convex mirror?

4. What is the focal length of a thin lens with one convex surface of 25-cm radius of curvature and a concave surface of 50-cm radius, the index of refraction of the material being 1.60?

5. Compute the focal length and the location of the cardinal points of the lens of problem 4 if the thickness of the lens is 2 cm. Indicate the locations of the principal points on a diagram.

6. What are the object and image locations for a converging lens of 25 diopters power, the magnification being 20 diameters?

7. A prescription calls for a $+4.00$ vertex diopter spectacle lens, and the lens blank, whose index is 1.523, has one finished surface of $+50$ -mm radius. If the thickness of the lens is 4 mm, what should be the radius of the concave surface?

8. A prescription calls for a -4.00 vertex diopter spectacle lens, and the lens blank, whose index is 1.523, has one finished surface of $+117$ -mm radius. If the thickness is 1 mm at the center, what should be the radius of the concave surface?

9. A thin positive lens of 30-cm focal length is combined with a thin negative lens of 50-cm focal length 20 cm behind it. If the object is 40 cm in front of the first lens, what is the location of the final image and what is its magnification?

10. Solve problem 9 if the lenses are both 1 cm thick, are equiconvex and equiconcave, respectively, and are made of glass having an index of refraction

of 1.52. The separation of 20 cm is measured between the second vertex of the first lens and the first vertex of the second lens, and the 40-cm object distance is measured from the first vertex.

11. Locate the cardinal points and determine the focal length of the lens combination described in problem 10.

12. A lens is made of glass having an index of refraction of 1.5171 for sodium light. Its thickness is 5 mm, and its radii are respectively 150 mm and 75 mm, both surfaces being convex.

(a) Locate the cardinal points and determine the focal length.

(b) Check the focal length by $f = f' = f_1'(-s_2'/s_2)$.

(c) Compute the optical power and the vertex power.

(d) Compute the magnification and location of the image of an object 20 cm from the first vertex, using:

(1) the step-by-step method,

(2) the cardinal points of the thick lens.

13. Two thin converging lenses of 4 diopters each are placed 10 cm apart on a common axis. What are the location and the size of the image of an object 5 cm tall at a distance of 15 cm from the first lens?

14. A camera lens has a focal length of 10 in. What is the elevation at which it will just photograph 1 square mile of the earth's surface on a 5 x 5 in. plate?

15. A biconvex lens of index 1.54 has a focal length of 40 cm in air. What will be its focal length in water, $n = 1.33$?

16. What is the shortest possible distance between an object and its real image in the case of a converging lens? Show that, if the distance is greater than this amount, two positions of the lens may be found between the object and its image, and that, if I_1 and I_2 are the sizes of the image in these two cases, the size of the object may be found by $O = \sqrt{I_1 I_2}$.

17. An object 5 cm high is located 30 cm from the convex vertex of a concavo-convex (converging) lens of index 1.50 whose radii of curvature are respectively 5 cm and 10 cm, the vertices being 10 cm apart. Determine the location, nature, and size of the image.

18. Locate the cardinal points of the lens specified in problem 17. Indicate their location on a diagram.

19. An object is located 20 cm from the surface of a sphere of glass whose index of refraction is 1.50 and whose radius of curvature is 6 cm. Determine the location, nature, and magnification of the image formed by paraxial rays.

20. A telephoto lens combination is made of a thin converging lens whose focal length is 4 cm and a thin diverging lens whose focal length is -2.5 cm mounted 2.5 cm behind the first on a common axis. Determine the focal length of this lens, and locate its cardinal points.

21. A converging meniscus lens is made of material whose index of refraction is 1.50, the radii of curvature being 100 mm and 50 mm with an axial thickness of 10 mm. What are the location and the size of the image of an object 5 cm high and 20 cm from the convex surface?

22. A positive meniscus spectacle lens is made of glass having an index of refraction of 1.523. If the surfaces have powers of $+12.00$ and -5.00 diop-

ers, respectively, and the vertex thickness is 4 mm, (a) what are the locations of the cardinal points? (b) What is the vertex power of the lens? (c) What are the radii of curvature of the two surfaces?

23. Derive a formula for the focal length of a cemented achromat in terms of the radii $r_1 = -r_2 = -r_3 = r$, and $r_4 = \infty$, the thicknesses t_1 and t_2 , and the indices of refraction n_1 and n_2 .

24. What must be the thickness and the radius of curvature of the convex surface of a plano-convex lens whose optical powers and vertex powers are, respectively, equal to those of a biconvex lens for which $r_1 = -r_2 = 100$ mm and $t = 10$ mm? Let $n = 1.50$ for both lenses.

25. Locate the cardinal points of a lens combination made up of a converging lens of 12.5-cm focal length and 9-mm thickness and a diverging lens of 16.5-cm focal length and 3-mm thickness, the vertices being 3 cm apart.

Chapter 4. Physical Limitations of Optical Images

4.1 General Considerations

Thus far the physical problem of image formation has been limited to rays in the vicinity of the principal axis and to ray pencils of very small aperture. In this chapter we will consider the conditions under which it may be possible to produce images of object points far removed from the principal axis, and the conditions under which the system may have a large aperture. A discussion of (1) field of view, (2) aperture, and (3) definition will show that it is physically impossible to extend all three simultaneously to any great degree. Compromises must be made according to the purpose of the optical system.

A plane mirror is unique in that it imposes no insuperable restrictions on field, aperture, and definition. Aplanatic surfaces do not restrict (2) and (3), but the useful field of view is limited to infinitesimal areas around the aplanatic points unless Abbe's sine condition is satisfied. This condition is important in the design of microscope and telescope objectives. The picturing of large areas by narrow beams is exemplified by the eyepieces of optical instruments and by conventional photographic objectives. In the latter it is possible to obtain large aperture with some sacrifice of definition or field of view or both. Because of the grainy nature of the photographic emulsion, it is in practice not necessary to strive for the highest possible resolution.

4.2 Abbe's Sine Condition

The sine condition of Abbe shows clearly the limitations of actual optical systems as compared with the ideal. The formula may be derived in various ways. Helmholtz and Clausius deduced it by using the principle of conservation of energy in the theory of radiation. The method used here is a simplified derivation first given by Hockin. In Fig. 4.1 VS_1 and $V'S_1'$ represent the first and last surfaces of an aplanatic system. Optical paths from

object points to image points will be represented by quantities in parentheses. Thus, since the system is aplanatic, we have by Fermat's principle:

$$(PVP') = (PSS'P')$$

Let P_1 be an object point a small distance above P , and let P_1' be its conjugate. For the paraxial ray paths

$$(PVP') = (P_1AF'P_1')$$

Assume that the system is aplanatic with respect to P_1 and P_1' so that

$$(P_1AF'P_1') = (P_1S_1S_1'P_1')$$

It is desired to find the condition under which this is possible, so that an optical system may image points off the axis as well as

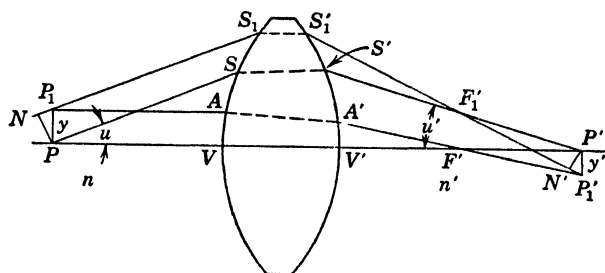


FIG. 4.1. Hockin's proof of Abbe's sine condition. ✓

points on the axis. If the above three equations are true, it must follow that

$$(P_1S_1S_1'P_1') = (PSS'P')$$

By using the normals PN and $P'N'$, this may be put into the form

$$-(NP_1) + (NP_1S_1S_1'N') + (N'P_1') = (PSS'P')$$

Now,

$$(NP_1S_1S_1'N') = (PSS'P')$$

so that

$$(NP_1) = (N'P_1')$$

By introducing the distances $y = PP_1$ and $y' = P'P_1'$, and the slope angles u and u' , this becomes

$$ny \sin u = n'y' \sin u'$$

which is Abbe's sine condition in its most general form. All good physical lens systems satisfy this condition with more or less accuracy. Since it differs from Lagrange's law for ideal optical systems, it must be concluded that actual systems can at best only approximate ideal optical systems.

4.3 Aberrations

The deviations of actual systems from ideal optical systems are called aberrations. They are (a) spherical aberration, (b) coma, (c) astigmatism, (d) curvature of image, (e) distortion, (f) longitudinal chromatism, (g) lateral chromatism, and (h) chromatic variation of spherical aberration. These will be defined and discussed in turn.

Spherical aberration is the change in location of the image of a point on the principal axis produced by rays through different annular zones of the lens system. **Coma** is a similar defect, except that the object point is off the axis and the image is a one-sided blur. **Astigmatism** is the formation of two mutually perpendicular line images of any object point. If the system is axially symmetrical, there will be no astigmatism unless the object is off the principal axis. **Curvature of image** means the formation of an image of a plane object on a curved surface in the image space. **Distortion** is the imaging of straight lines as curves. It is primarily due to a variation of magnifying power with object distance from the principal axis. In **longitudinal chromatism** the images produced by light of different colors are at different distances along the axis, while in **lateral chromatism** the images have different magnifications. Since magnification depends primarily on focal length, assuming a distant object, lateral chromatism refers to a variation in focal length with wave length. Under similar conditions, longitudinal chromatism refers to a variation in the location of the principal focus with wave length. If the lens system is not thin, these are distinct aberrations, and the elimination of one does not eliminate the other.

Logically, each of the first five "monochromatic" aberrations depends on wave length, but in all cases except the first, the variation is practically always negligible. They are called **monochromatic aberrations** because they are present even if a single wave length is used.

Von Seidel, in 1855, defined all these aberrations quantitatively and derived formulas for them, using the first two terms of the series expansion of the sine function:

$$\sin u = u - \frac{u^3}{3!}$$

Consequently his formulas are said to give the magnitude of the aberrations of the "third order." In the case of spherical aberration, formulas have been developed based on the expansion of the sine function to higher orders. However, the formulas become so complex, in general, that it is preferable to calculate ray paths trigonometrically when investigating this aberration. If the work sheet is properly laid out,¹ this is not a difficult task and has the advantage that no approximations are required. The rays which do not intersect the principal axis, called skew rays, are much more difficult to compute and are consequently often ignored.

The third order aberrations are discussed below in order to illustrate their nature better and to show how they may be corrected. As a rule, derivations of the formulas will be omitted. They may be found in treatises on optical design.

4.4 Spherical Aberration

The change in image location s' with radius h of the lens zone is the **longitudinal spherical aberration**. However, the change in $1/s'$ gives a simpler formula, which has been expressed in a convenient form by Coddington by the use of simplifying factors defined as follows:

(a) Shape factor

$$u \equiv \frac{R_2 + R_1}{R_2 - R_1}$$

so that

$$\frac{R_1}{R_2} = \frac{u - 1}{u + 1}$$

(b) Position factor

$$v \equiv \frac{s' - s}{s' + s} = \frac{2f}{s} - 1 = 1 - \frac{2f}{s'}$$

¹ For one example, see HARDY and PERRIN, *The Principles of Optics*, McGraw-Hill, 1932, page 38.

For example, position factors of -1 , 0 , and $+1$ correspond respectively to object distances of infinity, $2f$, and f . For positions on the axis, Coddington finds for a simple thin lens that

$$\Delta\left(\frac{1}{s'}\right) = \frac{h^2}{f^3} (Au^2 + Buv + Cv^2 + D) \equiv \frac{h^2}{f^3} S$$

where

$$A = \frac{n+2}{8n(n-1)^2}$$

$$B = \frac{n+1}{2n(n-1)}$$

$$C = \frac{3n+2}{8n}$$

$$D = \frac{n^2}{8(n-1)^2}$$

The change in s' , giving the longitudinal spherical aberration, may be found by the use of

$$\Delta s' = -s'^2 \Delta\left(\frac{1}{s'}\right)$$

Since the above quadratic equation for $\Delta(1/s')$ has no real roots for $n > 0.25$, spherical aberration cannot be zero for any simple lens. It can, however, be minimized as is shown below. To eliminate it entirely, a combination of at least two lenses must be employed.

The best shape for a simple thin lens is found by setting

$$\frac{\delta\left(\Delta\frac{1}{s'}\right)}{\delta u} = \frac{h^2}{f^3} (2Au + Bv) = 0$$

giving

$$u = -\frac{B}{2A} v$$

For a distant object, $v = -1$; hence $u = B/2A$ or

$$u = \frac{(n+1)8n(n-1)^2}{2n(n-1)2(n+2)} = 2\left(\frac{n^2-1}{n+2}\right)$$

which gives the condition for minimum spherical aberration in this special case. If, moreover, $n = \frac{3}{2}$, one finds that $u = \frac{5}{7}$ and that $R_1/R_2 = -\frac{1}{6}$, making $R_2 = -6R_1$. A converging and a diverging lens having this relationship of their radii are shown in Fig. 4.2. For other indices and object locations, the best shapes will vary decidedly from these illustrations. Distant images would call for a reciprocal ratio of the radii.

It should be noted that this discussion ignores all other aberrations and that the simple lens giving the least spherical aberration may be expected to have increased aberrations of other kinds. The meniscus lens on cheap cameras is an excellent illustration of a lens that is deliberately shaped or turned "incorrectly" as far as spherical aberration is concerned. This is done

to obtain a flat image, which is more important than some S.A. in a photographic lens. The comparatively large S.A. is reduced by using the lens at relatively low apertures of $f/11$ or so.

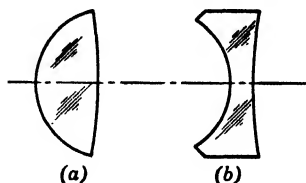


FIG. 4.2. Simple lenses having minimum spherical aberration for a distant object at the left, the lenses having an index of refraction of 1.5. (a) Converging. (b) Diverging.

4.5 Coma

This zonal aberration for oblique rays is difficult to observe in practice because it is usually mixed with a relatively large amount of astigmatism. The nature of the aberration is illustrated in Fig. 4.3, which shows a zone (a) of the lens L and the comatic circle (c) formed by rays passing through that zone. Rays through points 1, 5 of the zone focus at point 1 on the circle, rays through 3, 7 focus at point 3 on the comatic circle, etc. A simultaneous displacement of the centers of the comatic circles, Fig. 4.3(b), gives rise to a comet-shaped appearance of the focus. The "coma" is said to be positive if the tail of the "comet" extends away from the axis, as in the figure. The diameter of each comatic circle is proportional to the image height and to the square of the radius of the lens zone.

Abbe's sine condition must be satisfied if an optical system is to be free of coma. This requires that the transverse magnification y'/y be constant for all zones. Thus the general formula $n'y' \sin u' = ny \sin u$ is changed to $\sin u'/\sin u = \text{constant}$, for the

usual case where $n' = n$. In many cases, e.g. telescope and camera lens, the object is very far away, so that $\sin u$ is proportional to h . Then the sine condition reduces to

$$\frac{h}{\sin u'} = \text{constant}$$

which is easily tested from the slope angles u' obtained by tracing rays trigonometrically through the optical system.

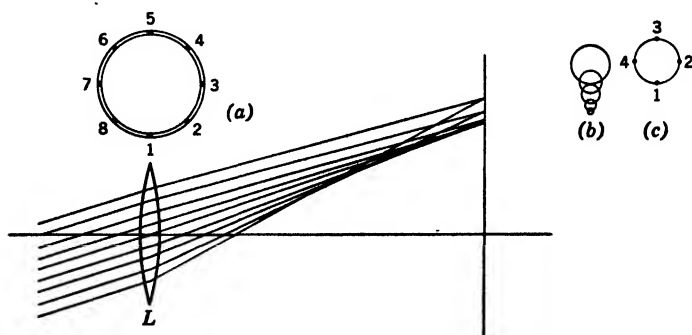


FIG. 4.3. Formation of comatic circles by the zones of a lens. (a) Points on a zone of the lens. (b) A family of overlapping comatic circles is produced by several zones. (c) Foci of diametrically opposite rays form a comatic circle.

Referring to Fig. 4.4, which shows conjugate rays, one observes that their intersection P' should be in the second principal plane, that I' is the image location for a distant object, and that $P'I' =$

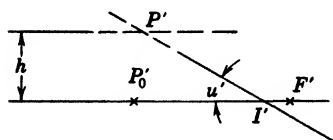


FIG. 4.4. The sine condition requires that $P'I'$ be constant. The distance $I'F'$ indicates the amount of longitudinal spherical aberration, L.S.A.

$h/\sin u'$. If there is no spherical aberration, I' is at the same place for all values of h . Therefore the condition that $h/\sin u' = \text{constant}$ defines a sphere about I' as the locus of points P' , instead of a plane as in the ideal system or in the first order theory. This requirement shows again that actual physical systems deviate from the ideal whenever it is not legitimate

to approximate $\sin u$ by u and $\sin u'$ by u' .

If longitudinal spherical aberration of amount L.S.A. is present, the sine law $h/\sin u' = \text{constant}$ does not indicate the absence of

coma. In this case, one must take $f_0' - \text{L.S.A.}$ as the distance from I' to the principal point P_0' on the axis, the subscripts indicating values of the usual quantities when h approaches zero. Thus $P_0'I'$ should equal $h/\sin u'$ if no coma is to be present when there is spherical aberration. Consequently, the difference

$$(f_0' - \text{L.S.A.}) - \frac{h}{\sin u'} = \left(f_0' - \frac{h}{\sin u'} \right) - \text{L.S.A.}$$

is taken as a measure of the amount of coma, while the quantity

$$f_0' - \frac{h}{\sin u'}$$

is called the deviation from the sine condition. These quantities are readily obtained by trigonometrical computation. Typical results are plotted for a series of simple lenses in some very instructive graphs in Hardy and Perrin's *The Principles of Optics*, page 88. The study of the change of these aberrations as the shape of the lens is altered, while its focal length is kept constant, is called **bending the lens**.

4.6 Distortion

A pinhole image is not distorted, as one can readily prove geometrically. To obtain a similar geometrical construction in the case of a lens system, the aperture stop (Art. 5.1) is first replaced

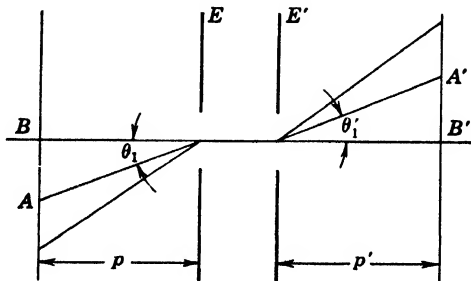


FIG. 4.5. Distortion related to slope angles of chief rays.

by its images in the object space and in the image space. These are, respectively, called the **entrance pupil** and the **exit pupil**, and they serve to limit the apertures of the cones of rays entering and leaving the optical system. The rays passing through the

centers of these pupils are the **chief rays** which are drawn in Fig. 4.5. Let p and p' denote the distances between the conjugate planes and the corresponding pupils E and E' . It is readily seen, Fig. 4.5, that the lateral magnification is given by

$$\beta \equiv \frac{A'B'}{AB} = \frac{p' \tan \theta_1'}{p \tan \theta_1}$$

When β is independent of θ , there is no distortion.

If it is true that p'/p is independent of θ , the condition for no distortion may be simplified to the statement

$$\frac{\tan \theta'}{\tan \theta} = \text{constant}$$

The constancy of this ratio may be readily checked by trigonometrical ray tracing. For a symmetrical doublet with the aperture stop centrally located between the two components, $\theta = \theta'$ because of symmetry, and the ratio

$$\frac{\tan \theta'}{\tan \theta} = 1$$

Such a lens is usually called a **rectilinear** or **orthoscopic lens**. However, the ratio p'/p is generally not independent of θ unless the lens components are free from spherical aberration with respect to the location of the aperture stop. If this is not the case, the change in the ratio p'/p must be taken into consideration by retaining this term in the original formula.

Even when the components do not satisfy the condition of constant p'/p , a symmetrical doublet is still orthoscopic for a unit linear magnification, since the changes in p and p' with θ are equal because of symmetry. Then $p' = p$ and $p'/p = 1$, a constant, for all angles.

It is evident that the location of the aperture stop and the residual spherical aberration are important in the correction of a lens system for distortion. It also follows that a lens giving no distortion in copying or photoengraving may give considerable distortion if used on distant objects as in aerial photography or surveying.

Distortion is classified as positive if the magnification increases off the axis, so that a square object has a pincushion-shaped image.

The converse, leading to a barrel-shaped image of a square, is called negative distortion. With a single lens, positive distortion is accentuated by a stop placed behind the lens, while negative distortion is increased if the stop is in front. The amount of distortion in per cent is given by

$$D = \frac{\text{zonal image height} - \text{paraxial image height}}{\text{paraxial image height}} \times 100$$

In this equation, magnifications may be used instead of image heights. Distortion varies about as the cube of the image height, but is independent of aperture.

4.7 Astigmatism

When light from an object point passes obliquely through a lens or lens system, the image is usually found to consist of two lines at right angles to each other and some distance apart. This is due to a different focussing power in different meridians. The emerging light wave (orthotomic surface) may be thought to have a shape similar to a patch on an automobile tire, having different principal radii of curvature in two meridians at right angles to each other. The focussing of such a wave is illustrated in Fig. 4.6. The rays in

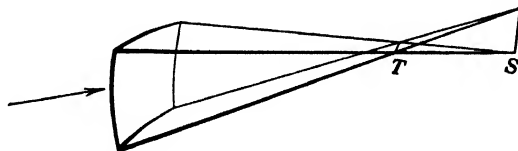


FIG. 4.6. An astigmatic pencil of rays, showing the two line foci at *T* and *S*.

planes of least curvature focus in a line at a greater distance than the rays in planes of greatest curvature. The shape of the figure outlined by such rays from a circular base on the wave front is called a **Sturm's conoid**.

If, instead of a point, the object consists of a star-shaped figure, the image will be made up of superposed line elements corresponding to the various object points. As shown in Fig. 4.7, the line elements will overlap for some direction of the arm of the cross, giving a sharp image for this line. There will be another astigmatic focus at another location in which the arm at right angles to the first will be in sharp focus. One or both of these astigmatic

foci may be virtual. In testing lenses for astigmatism, one often uses an object made up of radial lines and circles with their centers on the principal axis of the lens. The astigmatic image of the circular arcs is called the **primary** or the **tangential** focus, while the image of the radial lines is called the **secondary** or the **sagittal** focus. For a simple converging lens, the primary focus lies nearer the lens than the secondary, whereas for a simple diverging lens, the opposite is true.

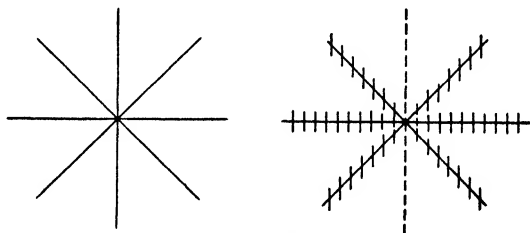


FIG. 4.7. Astigmatic image of a star-shaped object. One arm is in sharp focus.

For a single refracting surface, the distances along the central rays from the surface to an astigmatic focus are given by the equations

$$\frac{n \cos^2 i}{s} + \frac{n' \cos^2 i'}{s_1'} = \frac{n' \cos i' - n \cos i}{r}$$

for the primary focal distance s_1' , and

$$\frac{n}{s} + \frac{n'}{s_2'} = \frac{n' \cos i' - n \cos i}{r}$$

for the secondary focal distance s_2' . As usual, s is the object distance, i is the angle of incidence, i' is the angle of refraction, n and n' are respectively the indices of the object and image space, and r is the radius of curvature of the refracting surface.

For a spherical mirror, one may write $n = -n'$, since this convention changes Snell's law into the law of reflection.

For a thin lens with a central stop of small aperture, Coddington has shown that

$$\frac{1}{s} + \frac{1}{s_1'} = \frac{1}{\cos i} \left(\frac{1}{r_1} - \frac{1}{r_2} \right) \left(\frac{n \cos i'}{\cos i} - 1 \right)$$

and

$$\frac{1}{s} + \frac{1}{s_2'} = \cos i \left(\frac{1}{r_1} - \frac{1}{r_2} \right) \left(\frac{n \cos i'}{\cos i} - 1 \right)$$

where i and i' are respectively the angles of incidence and refraction at the first lens surface. Thus $\sin i = n \sin i'$, where n is the index of the lens.

One may readily prove that the astigmatic difference $s_2' - s_1'$, in the case of a thin lens, is

$$s_2' - s_1' = s_1' s_2' \left(\frac{1}{r_1} - \frac{1}{r_2} \right) \left(\frac{n \cos i'}{\cos i} - 1 \right) \sin i \tan i$$

while for a mirror it is

$$s_2' - s_1' = \frac{2s_1' s_2'}{r} \sin i \tan i$$

For a single lens or doublet, the successive order, in the direction of the light propagation, of the tangential focus T , the sagittal focus S , and the Petzval surface P (defined below) is TSP in the case of positive astigmatism, and PST in the case of negative astigmatism. The astigmatic difference of a two-lens combination may be altered by changing the curvatures. As the astigmatism is thus reduced, the image approaches a parabolic surface, called the **Petzval surface**. Once the powers of the lenses and their indices of refraction are determined, the form of the Petzval surface is fixed. If the combination is overcorrected, the order of TSP is reversed.

The amount of astigmatism, as measured by the difference in focus, varies as the square of the image height. For compound lenses, the Petzval surface loses its parabolic form, and in many anastigmats lies between T and S . In all cases, the axial distance PT is three times the axial distance PS .

As the thin-lens equations show, astigmatism varies only slightly with the shape of the lens, but it reverses sign with the power of the lens. It should be noted, however, that in spectacle lenses the passage of narrow pencils of rays through off-center portions of the lens introduces a so-called **marginal astigmatism** which varies decidedly with the shape of the lens. Modern meniscus lenses are designed to reduce this marginal astigmatism to a minimum. Such lenses are discussed in Art. 7.2 in connection with ophthalmic optics.

It is possible to design complex combinations of lenses having appropriate separations so that the resulting product is relatively free from astigmatism. These are widely used as photographic objectives, since astigmatism is particularly troublesome here in

connection with images projected some distance from the principal axis. The primary and secondary astigmatic foci of such anastigmats generally intertwine as shown in Fig. 4.8, which applies to a Zeiss Sonnar $f/2$ lens of 50-mm focal length. All high-grade photographic objectives are designed so that not only astig-

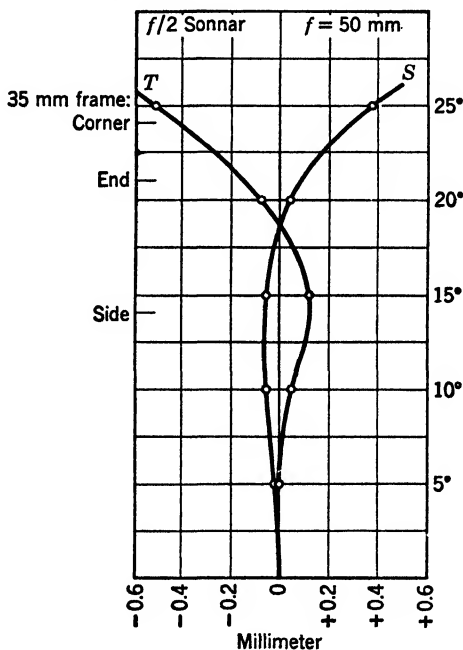


FIG. 4.8. Tangential and sagittal foci of a Zeiss Sonnar lens, illustrating the residual astigmatism in a typical anastigmatic lens. The ordinates are slope angles of chief rays, and the vertical line at 0 represents the focal plane.

matism, but all the aberrations discussed in this chapter, are reduced below a tolerable value for a certain object distance and field of view. The latter restrictions and the magnitude of the tolerances depend on the purpose for which the lens is designed. For general-purpose lenses, the object is assumed to be at "infinity," the field of view being about 52° . A resolving power of between 50 and 100 lines per millimeter is usually acceptable, since this exceeds the resolving power of most photographic emulsions. A classification of photographic objectives has been given by Kingslake in the *Journal of the Optical Society of America*,

12, 251, and in the *Handbook of Photography* edited by Henney and Dudley.

4.8 Curvature of Image

In the case of an anastigmat, the locus of the circle of least confusion, or best stigmatic image, lies between the two astigmatic images at about one fourth the distance from the sagittal image to the tangential image. If the astigmatism is not too great, a fairly sharp image is found here. The locus of the image points corresponding to points on a distant plane perpendicular to the principal axis defines a surface which indicates the curvature of the image by its deviation from a plane. This is approximately the Petzval surface for this case. In properly designed objectives it must approximate a plane within the useful field of the lens.

Petzval showed that the condition for a flat image by a doublet made of two thin lenses, separated or in contact, is

$$\frac{1}{f_1 n_1} + \frac{1}{f_2 n_2} = 0$$

4.9 Chromatic Aberration

Because the index of refraction of any transparent material varies with wave length, the focal length and locations of the cardinal points will also vary with the wave length. As usual, this aberration may be corrected, at least for two or three wave lengths, by the use of a suitable combination of lenses. In the case of thick lenses, chromatic aberration is classified as longitudinal or lateral, depending on whether the location or the size of the image varies with wave length. In general, these aberrations require separate corrections. Elimination of the longitudinal aberration requires essentially that the location of the principal focus be the same for two or three wave lengths, while the correction for the lateral aberration requires that the focal length be the same. In the case of a thin lens where the principal planes do not change their location appreciably with color, the correction of one chromatic aberration simultaneously eliminates the other.

For a thin lens $1/f = (n - 1)K$, where K represents a shape factor which is independent of wave length. The change in power due to a change Δn in index of refraction is given by

$$\Delta \left(\frac{1}{f} \right) = \Delta n \cdot K = \frac{\Delta n}{n - 1} \left(\frac{1}{f} \right)$$

If we select the usual C and F Fraunhofer lines as the wave lengths for which Δn is the change in index, we may write

$$\Delta \left(\frac{1}{f} \right) = \frac{\omega}{f}$$

where ω is the dispersive power

$$\omega = \frac{n_F - n_C}{n_D - 1}$$

Other pairs of lines or wave lengths may be chosen in special cases. The D line of sodium is, however, the customary standard wave length for the index n when its variation is not considered.

The above equation shows that the change in power is dependent on the kind of glass and the power of the lens, and that it reverses in sign with the power. Thus one may compensate the change in power of one lens by a change in power of equal and opposite amount in another lens combined with it. In general, one may write

$$\frac{1}{f} = \frac{1}{f_1} + \frac{1}{f_2} - \frac{d}{f_1 f_2}$$

Therefore,

$$\begin{aligned} \Delta \left(\frac{1}{f} \right) &= \frac{\omega_1}{f_1} + \frac{\omega_2}{f_2} - \left(\frac{d}{f_1} \right) \left(\frac{\omega_2}{f_2} \right) - \left(\frac{d}{f_2} \right) \left(\frac{\omega_1}{f_1} \right) \\ &= \frac{\omega_1}{f_1} + \frac{\omega_2}{f_2} - \frac{d(\omega_2 + \omega_1)}{f_1 f_2} \end{aligned}$$

The pair of lenses will be an achromatic combination if

$$\Delta \left(\frac{1}{f} \right) = 0$$

Two special cases will be considered. If we have two thin lenses in contact, $d = 0$, the condition for achromatism is

$$\frac{\omega_1}{f_1} + \frac{\omega_2}{f_2} = 0$$

This combined with the equation

$$\frac{1}{f_1} + \frac{1}{f_2} = \frac{1}{f}$$

enables one to calculate the focal lengths f_1 and f_2 of the two component lenses for any given dispersive powers ω_1 and ω_2 and for a given focal length f . Since the ω 's are always positive, the focal lengths f_1 and f_2 must be of opposite sign. Note that ω_1 must not be equal to ω_2 , for then the first equation requires that $f_1 = -f_2$, and the second equation shows that $f = \infty$. Hence a combination of two different materials, such as crown glass and flint glass, is chosen. For a converging doublet the diverging lens must have the longer focal length and therefore the greater dispersive power.

It is possible to achromatize with respect to focal length, using two lenses of the same kind of glass, if the interval d is not zero. In this case, $\omega_1 = \omega_2$ and the original equation for change in power becomes

$$\Delta\left(\frac{1}{f}\right) = 0 = \omega\left(\frac{1}{f_1} + \frac{1}{f_2}\right) - \frac{2d\omega}{f_1f_2}$$

This equation is satisfied if

$$d = \frac{1}{2}(f_1 + f_2)$$

This principle is frequently employed in the construction of eyepieces according to Huygens' or Ramsden's formula. The achromatism obtained is valid only for the transverse aberration, not for the longitudinal. However, this is usually sufficient for an eyepiece.

The condition for achromatism of a cemented doublet previously given ($d = 0$) is inconsistent with Petzval's condition for flatness of field unless the glass of greater index has the smaller dispersive power. The barium glasses developed at Jena at the close of the last century were the first to satisfy this requirement.

PROBLEMS

1. Determine the longitudinal spherical aberration, the deviation from the sine condition, the coma, and the chromatic aberration of the lens specified in problem 1, Chapter 3.

2. (a) What form of lens, as defined by the ratio of its radii, gives the least spherical aberration for an object distance of 50 cm and an image distance of 30 cm? Let $n = 1.50$. (b) What is the magnitude of the longitudinal spherical aberration of this lens for rays entering at a height of 20 cm?

3. Compute the radii of curvature of a condensing lens which will give minimum spherical aberration for an object located 10 cm from its first prin-

principal plane with the image at 15 cm from the second principal plane. Let $n = 1.50$.

4. Derive a formula for the magnitude of the spherical aberration (*a*) for an equiconvex lens with the object and image at equal distances on opposite sides of the lens, and (*b*) for a plano-convex lens with the object at infinity and the convex side toward the object.

5. A converging meniscus lens has a diameter of 5 cm, a negligible thickness, and an index of refraction of 1.50 with radii of curvature of 100 mm and 50 mm. If an object point is 20 cm away on a line making an angle of 30° with the principal axis, what are (*a*) the astigmatic difference? (*b*) the length of the primary astigmatic line image?

6. A thin achromatic doublet of 20-cm focal length is to be made of crown glass and flint glass whose dispersive powers are 0.0141 and 0.0232, respectively. What are the required focal lengths of the components?

7. When a lens is achromatized only for its focal length, the magnification is practically the same for all colors. Explain, pointing out restrictions.

Chapter 5. Apertures of Optical Systems

5.1 Diaphragms or Stops

Diaphragms or lens rims have one or more of three principal effects in optical systems: (1) restricting the solid angle filled by cones of rays, (2) restricting the range of locations of objects and images to the useful field of the system, and (3) reducing internal reflection. Restrictions (1) and (2) are usually effected with the idea of obtaining an economical design and construction, consistent with tolerable or desired performance. In the case of a camera lens one wishes to vary the aperture to obtain the desired depth of focus and perhaps to reduce the aberrations of a faulty lens.

The actual diaphragm or lens rim which limits the angular spread of the cones of rays is called the **aperture stop**. The extreme angle between incoming rays is called the **angular aperture**, while that between emerging rays is called the **angle of projection**. These angles are designated as $2U$ and $2U'$ respectively in Fig. 5.1, in which the diaphragm S is the aperture stop. A stop like that in Fig. 5.1(a), which is in the object space, is called a **front stop**. The stop shown in Fig. 5.1(b) is called a **back stop**.

5.2 Entrance Pupil and Exit Pupil

If the aperture stop is not in the object space, the size and location of its conjugate in object space may be found by the general formulas for optical systems. This conjugate then acts as a virtual stop which limits the aperture of the entering rays. That virtual or real stop in the object space which functions in this manner is called the entrance pupil. It is usually found by first transferring all diaphragms and lens rims into the object space by means of the conjugate focal equations. The stop or image of stop in the object space which is found to subtend the smallest angle at the object location is the **entrance pupil**, for example, E in Fig. 5.1. The entrance pupil may be different for different object locations, but in a well-designed system it is usually the

same. The angle subtended at the object by the entrance pupil is called the **angular aperture** of the system, for example, $2U$ in Fig. 5.1. The ratio of the diameter of the entrance pupil to the focal length of the system is called the **relative aperture**; its

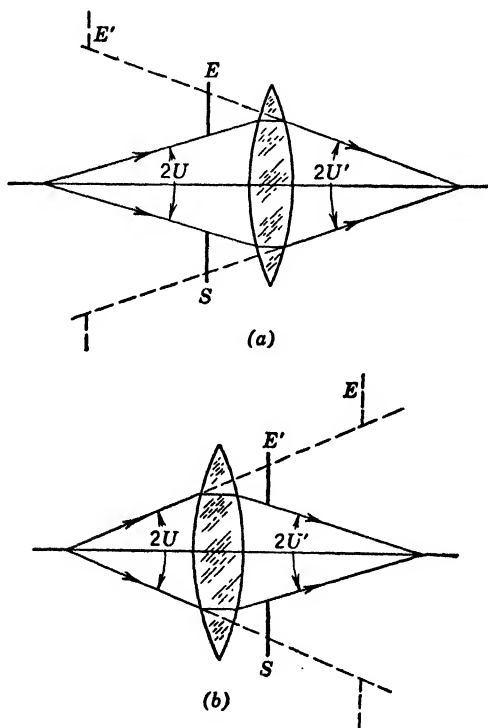


FIG. 5.1. (a) Locations of entrance and exit pupils of a lens with a front stop. (b) Similar diagram for a lens with a back stop.

reciprocal is the ***f*-number**. These quantities will be used later in discussing the light flux per unit area of image.

Similar considerations may be applied to the stops and images of stops in the image space. That one which subtends the smallest angle at the image is called the **exit pupil**, for example, E' in Fig. 5.1. The exit pupil is always an image of the entrance pupil with respect to the entire optical system, for it is the image of the aperture stop in the elements which follow the aperture stop, whereas the entrance pupil is the image of the same stop formed

by the elements which precede it. An application of the discussion in the preceding paragraphs to a problem such as number 4 at the end of this chapter will do much to fix these ideas in mind. Figure 5.2 shows how the entrance pupil may be found by transferring the lens rims of L_1 and L_2 and the stop S into the object space.

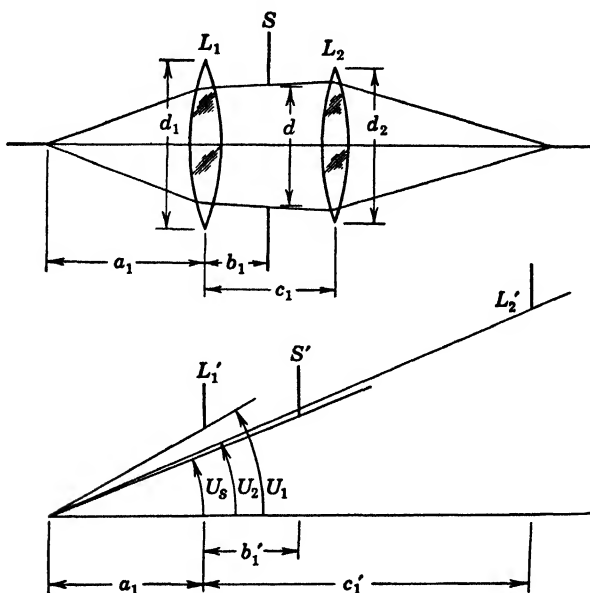


FIG. 5.2. Determination of the entrance pupil for a pair of lenses having a stop located between the components. The image S' of S is the entrance pupil, since it subtends the smallest angle at the object.

The smallest of the angles U determines which of the images is the entrance pupil.

5.3 Field of View

Any ray passing through the center of the entrance pupil, and therefore also through the center of the exit pupil, is called a **chief ray**. Such rays are axes of cones of rays which may enter the optical system. When the cones are made increasingly oblique, they will be cut off partially and finally completely by the other stops. This causes a restriction of the field of view of the optical system. That diaphragm or image of diaphragm in the object space which subtends the smallest angle from the center of the

entrance pupil is called the **entrance port**. The angle it subtends is called the **angle of the field of view**. The field of view will not be sharply bounded unless the entrance port is coincident with the object, since the entrance port cuts the entering cones of rays to their axes, thus limiting the entering light flux to about half the maximum. The image will be vignetted correspondingly unless the cones are cut at their apex, which is the case only if the entrance port lies on the object. The entrance pupil can never be the entrance port, since it subtends the largest possible angle, 180° , at its own center.

5.4 Depth of Field and Focus

Another important function of a diaphragm is the control of the depth of field and the depth of focus. These are, respectively, the variations in object distance and in image distance which can be tolerated without incurring an objectionable lack of sharpness in focus. If an optical system whose entrance and exit pupils are at E and E' , Fig. 5.3, forms an image P' of an object at P , another

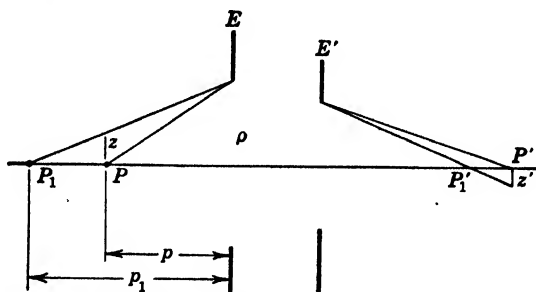


Fig. 5.3. Determination of the depth of field of a lens system.

object point at P_1 will be out of focus, producing a circle of confusion of radius z' at P' . If the object in critical focus is at a distance p from the entrance pupil, the problem is to find the distance p_1 at which objects will be out of focus by just the maximum tolerable amount as specified by a permissible radius of the circle of confusion z' . From the geometry of Fig. 5.3, it follows that

$$\frac{p_1 - p}{z} = \frac{p_1}{\rho}$$

where z is the conjugate of z' in the object space, and ρ is the radius

of the entrance pupil. Using the formula for transverse magnification

$$z' = \frac{f}{x} z$$

to eliminate z , one readily finds that

$$p_1 = \frac{p}{1 - \frac{z'x}{f\rho}}$$

The quantity $f\rho/z'$ is called the **hyperfocal distance**, and may be designated by the symbol x_0 . The formula then becomes

$$p_1 = \frac{p}{1 - \frac{x}{x_0}}$$

Similarly one may show that objects closer than P are in tolerably good focus down to a distance

$$p_2 = \frac{p}{1 + \frac{x}{x_0}}$$

As to the physical significance of the hyperfocal distance, it should be noted that, if $x = x_0$, then $p_1 = \infty$, and $p_2 = \frac{1}{2}p$. Thus for the special case of objects in critical focus at a distance equal to the hyperfocal distance, the depth of field extends from $\frac{1}{2}p$ to infinity. Since the value of p is very nearly equal to $f + x$, the value of x for distant objects does not differ greatly from p , especially for a lens of short focal length.

The magnitude of the hyperfocal distance, and therefore the depth of field, depends on the permissible value of z' as well as on f and ρ . In photography, one usually sets $z' = 0.002$ in., while ρ is obtained from the relative aperture $2\rho/f$ of the objective and its focal length f . It should be noted, however, that the computed results p_1 and p_2 are meaningless if a z' is assumed which is smaller than the lens is capable of giving because of its aberrations, because of diffraction, or because of graininess of the emulsion. If the lens is imperfect, for example, one must use an appropriately

larger z' which leads to a greater depth of field, as it would with any lens. The only difference is that the quality of the focus with the poorer lens will be more uniformly poor over the entire range from p_1 to p_2 .

5.5 Perspective

A three-dimensional object whose vertical section is $ABCD$, Fig. 5.4, will be imaged as a three-dimensional image like $A'B'C'D'$, the points N and N' in the figure being the nodal points of the lens system. Because of depth of focus, the image may be projected

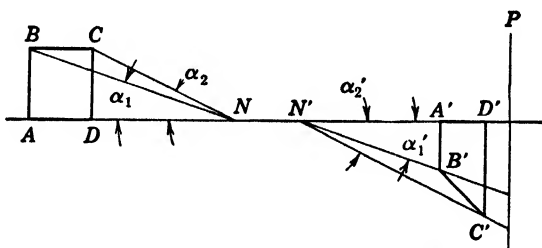


Fig. 5.4. Projection of an image onto an image plane at P .

on some plane such as P , Fig. 5.4, and observed or photographed there as a two-dimensional image. If one is to obtain the same perspective as with a monocular view of the object from N , one must view the image from N' , for in that case all angles α' will be the same as the conjugate object space angles α . Hence the image on the retina will be like that obtained by direct viewing of the object from N , and the perspective will then be the same. At greater viewing distances of the image than $N'P$, the angles α' are equal to those for a longer object, as one can see by considering B to be moved backward. In this case the perspective is increased. The proper viewing distance p' is approximately equal to the focal length of the lens used in forming the image. Since this focal length is often less than the reading distance of 25 cm or 10 in., especially when a miniature camera is used, one cannot obtain correct perspective without some optical aid or by enlargement of the image. If the image at P , Fig. 5.4, is enlarged M diameters, the proper viewing distance will be M times as great as p' , since this will preserve the original angle relationships. One reason for the advisability of enlarging most negatives, especially those ob-

tained with a miniature camera, is the rectification of perspective when the print is viewed from a convenient distance. Equally important is the choice of a suitable "taking" distance which gives a pleasing perspective in the first place.

PROBLEMS

1. How long should a vertical mirror be in order that a person 5 ft 10 in. tall may just be able to see his full-length image in it? How high should it be placed if the eyes are 5 ft 4 in. from the floor?

2. How far above a magnifying lens of 5-cm focal length and 6-cm diameter should the eye be placed so that the observer may see the whole of an object 10 cm in diameter, the image being at infinity?

3. A camera lens consists of two thin lenses whose focal lengths are 20 cm and diameters are 4 cm. The lenses are mounted 4 cm apart with a stop of 3-cm diameter placed half-way between them. What are the f number and the relative aperture of this system?

4. Two converging lenses whose focal lengths are 12 cm and 24 cm respectively and whose diameters are 3 cm and 2 cm respectively are placed 8 cm apart on a common axis. A diaphragm whose diameter is 2.5 cm is placed midway between the lenses. (a) Determine the entrance and exit pupils of this combination for an object distance of 30 cm from the first lens. (b) What is the relative aperture? (c) Determine the location and size of the entrance port and of the exit port.

5. A telescope has for its objective a thin positive lens of 20-cm focal length and 5-cm aperture, and for its ocular a thin negative lens of 4-cm focal length and 2-cm aperture. (a) Determine the locations and sizes of the entrance and exit pupils, showing these on a diagram. (b) Determine the angular field of view in the object space and in the image space.

6. A telescope is constructed of two thin converging lenses whose focal lengths are 180 mm and 15 mm respectively. They are 195 mm apart, and a 7-mm diaphragm is placed in the focal plane of the first lens. The diameters of the objective and ocular are respectively 20 mm and 8 mm, and the entrance pupil of the observer's eye is 4 mm in diameter and is located at the exit pupil of the telescope. (a) What are the location and size of the entrance pupil for a distant object? (b) What are the location and size of the exit pupil of the telescope? (c) What is the field stop? (d) What is the angular field of view in the object space?

7. In the Galilean telescope (opera glass) the locations and sizes of the entrance and exit ports are dependent on the relative sizes of the exit and eye pupils. Show by diagram how the field of view is affected thereby. Refer to problem 5.

8. A portrait is taken with a camera focussed on a subject 3.5 ft from the entrance pupil of the objective. (a) If the objective has a focal length of 2 in., how much must the picture be enlarged to show true perspective when it is viewed from the normal reading distance of 10 in.? (b) What degree of en-

largement of the same negative is needed if the perspective is to correspond approximately to that obtained with an original taking distance of 7 ft? Why is it impossible to obtain exactly the same perspective in the two cases?

9. Compute the depth of field for a photographic objective of 5-cm focal length at $f/4$ relative aperture if the tolerable circle of confusion has a radius of 0.02 mm and the camera is focussed critically for an object distance of 6 meters?

Chapter 6. Photometry and Image Brightness

6.1 Photometric Quantities

Radiant flux may be expressed in ergs per second, watts, or microwatts emitted by a source of radiation. The radiation is usually not homogeneous, but distributed throughout the spectrum in accordance with some distribution curve as shown in Fig. 6.1. The ordinates are watts per unit wave length interval, P_λ , so that the energy radiated per second between any two wave lengths λ_1 and λ_2 is expressed by an integral

$$P = \int_{\lambda_1}^{\lambda_2} P_\lambda d\lambda$$

At a wave length of 5550 Å each watt of radiation corresponds to a light flux of 685 lumens (*defined below*). At other wave lengths the number of lumens per watt is proportional to the visibility

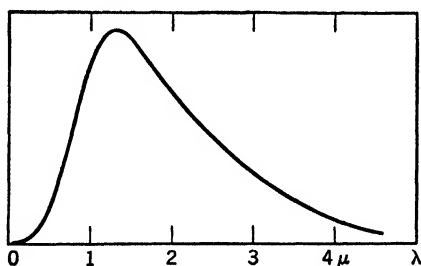


FIG. 6.1. Energy distribution as a function of wave length for a typical incandescent light source.

function V_λ , given in the Appendix, Table 6. Thus the total light flux in lumens is given by

$$F = 685 \int_0^\infty V_\lambda P_\lambda d\lambda$$

The **lumen** was formerly defined as the light flux from a uniform point source of one candle power into a unit solid angle, or into

one square foot at one-foot distance. The integral formula now relates this term to the concepts of spectro-radiometry instead of to a non-existent source of indefinite spectral distribution.

The **intensity of a source** will be defined as equal to the number of lumens radiated into a unit of solid angle, or better, lumens per unit solid angle. The magnitude of a solid angle ω is equal to the area intercepted by a cone on a unit sphere with its center at the vertex of the cone. Thus

$$I \equiv \frac{dF}{d\omega}$$

The units of I are *candle power*, as may be inferred from the old definition of the lumen. In fact, the factor 685 was originally derived from experiments comparing the two ideas of light flux, the physical and the psycho-physical. The flux in all directions from a source of intensity I is

$$F = \int_0^{4\pi} I d\omega$$

If the source radiates uniformly in all directions

$$F = I \int d\omega = 4\pi I$$

If the source is not uniform, one may define a mean spherical candle power, m.s.c.p., as

$$I_{\text{m.s.c.p.}} = \frac{F}{4\pi}$$

Illuminance is defined as the incident light flux per unit area

$$E \equiv \frac{dF}{da}$$

If the element of area da is illuminated by a relatively small, or "point," source at a distance r and if the angle of incidence is θ , the light flux dF reaching the area will be contained in a solid angle

$$d\omega = \frac{da \cos \theta}{r^2}$$

as can be seen in Fig. 6.2.

By the definition of I , the light flux

$$dF = I d\omega = I \frac{da \cos \theta}{r^2}$$

Therefore, the illuminance of da is given by

$$E \equiv \frac{dF}{da} = \frac{I \cos \theta}{r^2}$$

This is the well-known inverse square law with the cosine factor included. If I is in candle power and r is in feet, E is in **foot-**

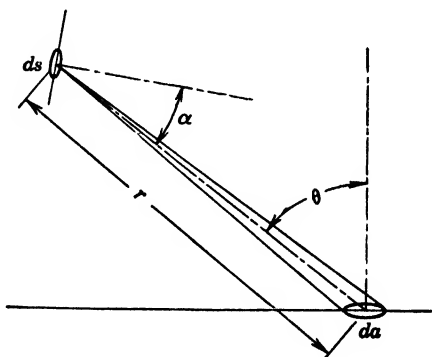


FIG. 6.2. Pencil of rays, from an element of area ds of the light source, which is obliquely incident on an element of area da . The cone subtends a solid angle $d\omega$.

candles, one foot-candle being identical with one lumen of incident light flux per square foot. If r is in meters, E is in **meter-candles**, or **lux**; while if r is in centimeters, E is in **phots**.

Although one may consider any light source as practically a point source at distances of many diameters, it is more accurate to treat the source as a luminous surface. In this case, its brightness or **luminance** is defined as the candle power per unit projected area. By projected area is meant the area of the source as seen from the illuminated element of area da . If the line joining da with the element of emitting surface ds makes an angle α with the normal to ds , the candle power of ds is

$$dI \equiv B ds \cos \alpha$$

where B is the luminance by definition.

For perfectly diffusing surfaces, B is independent of the direction of emitted or diffusely reflected light, so that the candle power varies as the cosine of the angle α . In such a case, Lambert's cosine law is said to hold. However, this is true only for ideal, not for actual, surfaces. Usually B falls off rapidly at angles of emission approaching 90° . In a few cases, notably that of glowing tungsten, the luminance B increases at first before falling rapidly as grazing emission is approached. It may be proved that, when Lambert's law is obeyed, a glowing sphere, or uniformly illuminated sphere, will have the appearance of a uniformly illuminated disk. Conversely, Lambert's cosine law is sometimes derived by assuming this identical appearance to be an experimental fact.

6.2 Brightness of Images

The photometric principles discussed above may be readily applied to determine the luminance and illuminance of optical images. Let ds be an element of area of the object, and let E , Fig. 6.3, be

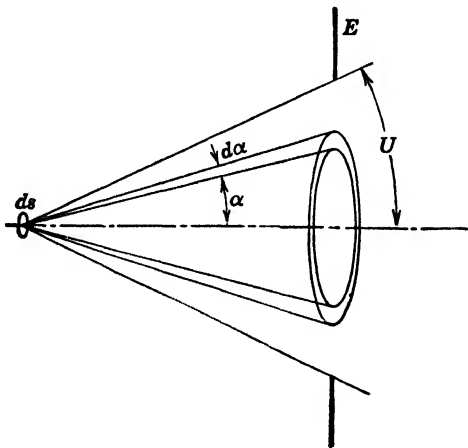


FIG. 6.3. Light flux from ds into the entrance pupil E of an optical system.

the entrance pupil of the optical system. The flux into the element of solid angle $d\omega$ defined by the double cone of aperture α to $\alpha + d\alpha$ is

$$dF = B ds \cos \alpha d\omega$$

Since the solid angle $d\omega$ is equal to the area of a unit sphere contained between the two cones

$$d\omega = 2\pi \sin \alpha \, d\alpha$$

Therefore

$$dF = 2\pi B \, ds \, d\alpha \cos \alpha \sin \alpha$$

The total light flux into the optical instrument is given by the integral

$$F_U = 2\pi B \, ds \int_0^U \sin \alpha \cos \alpha \, d\alpha = \pi B \, ds \sin^2 U$$

Similar considerations hold for the flux $F_{U'}$ passing through the

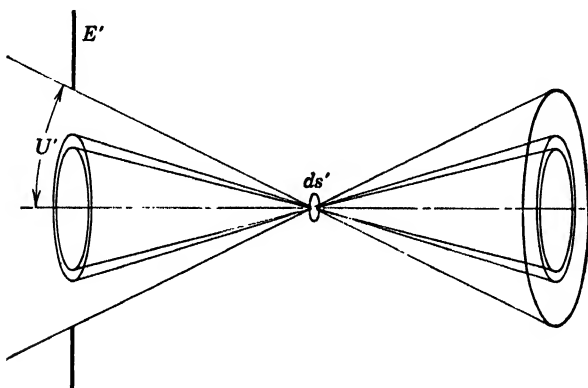


FIG. 6.4. Light flux emerging from the exit pupil E' of an optical system and passing through an element of area ds' of the image.

element of area ds' of the image which is conjugate to ds , Fig. 6.4. Hence

$$F_{U'} = \pi B' \, ds' \sin^2 U'$$

where B' is the luminance of the image, and $2U'$ is the angle of projection. In an ideal system in which there are no light losses $F_{U'} = F_U$, so that

$$B' \, ds' \sin^2 U' = B \, ds \sin^2 U$$

Now, Abbe's sine condition requires that

$$n y \sin U = n' y' \sin U'$$

in a coma-free system. Then

$$n^2 \, ds \sin^2 U = n'^2 \, ds' \sin^2 U'$$

and therefore

$$B' = B \left(\frac{n'}{n} \right)^2$$

This gives the greatest image luminance that can be obtained with an optical instrument of any possible design. Frequently $n' = n$; in that case the image luminance is at best just equal to the object luminance.

6.3 Illuminance of Images

In many applications, the light flux density or illuminance of the image is of greater importance than the luminance. Note that these are distinctly different quantities. If a given light flux is projected on a smaller area, the illuminance is definitely greater, but if at the same time the light flux is contained in a solid angle which is proportionately greater as the area is smaller, the luminance will be the same. In exposing a photographic plate, for example, the *illuminance* of the plate area is of prime importance, since the size of the angle of projection is immaterial. On the other hand, in illuminating a second optical system of fixed aperture, for example, the eye, the image *luminance* will determine the light flux into the optical system.

Image illuminance is given by the expression

$$E' = \frac{dF'}{ds'} = \pi B' \sin^2 U'$$

According to Abbe's sine condition,

$$\sin^2 U' = \frac{(n \sin U)^2}{n'^2 \beta^2}$$

where $n \sin U$ is called the **numerical aperture** of the optical system. Substituting the second expression into the first yields the result

$$E' = \frac{\pi B' (n \sin U)^2}{n'^2 \beta^2}$$

In most cases $n' = 1$, since the final image is in air. In the case of an image created by a photomicrographic system with an immersion lens, the luminance of the image ideally equals the lu-

minance B of the condenser diaphragm which is in air, as is the final image. Hence, for systems like the microscope,

$$E' = \frac{\pi B (n \sin U)^2}{\delta^2}$$

The illuminance is consequently proportional to the square of the numerical aperture and inversely proportional to the square of the linear magnification. Over a limited range, an increase in one can be compensated by an increase in the other.

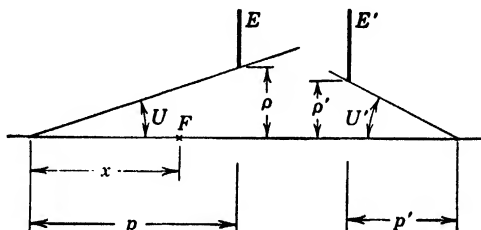


FIG. 6.5. Quantities pertaining to the flux density or illuminance of the image.

In systems like photographic lenses and telescopes, $n = n' = 1$. Moreover, the numerical aperture may be replaced by a simpler expression, as follows. From Fig. 6.5, it is apparent that

$$\tan U = \frac{\rho}{x}$$

If the object is so far away that the angle U is quite small, it is permissible to replace the tangent by the approximation

$$\sin U \doteq \frac{\rho}{p} \doteq \frac{\rho}{x}$$

where ρ is the radius of the entrance pupil, and x is the extrafocal distance of the object. Since the linear magnification is numerically equal to f/x , the image illuminance is expressible by the formula

$$E' \doteq \pi B \frac{\rho^2 x^2}{x^2 f^2} = \frac{\pi B}{4} \left(\frac{2\rho}{f} \right)^2 = \frac{\pi B}{4N^2}$$

where N denotes the **f-number** of the lens system and is equal to the focal length divided by the diameter of the entrance pupil.

The **relative aperture** of the lens or stop opening is generally written as f/N . In so far as the reciprocity law of photography holds, the time of exposure required for a given photographic density is proportional to the square of the f -number, N , a familiar photographic rule. In any actual optical system, the illuminance of the image area is also proportional to the transmissivity of the system.

It is easy to show, by omitting the approximation $p \doteq x$, that, in close-up photography or copying, the formula just given for the illuminance should also be multiplied by the factor x^2/p^2 . Thus, for 1:1 copying, $x = f$ and $p \doteq 2f$, so that the illumination is one fourth as large as for distant objects and the time of exposure must accordingly be four times as great.

6.4 Normal Brightness

The natural or subjective brightness of an object or image seen by the eye depends on the illuminance H_0 of the retina. This is given by the expression

$$H_0 = \pi B \sin^2 V$$

B being the luminance of the external object or instrument image, n being the index of refraction of the "vitreous humor," and V

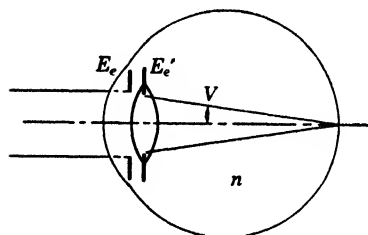


FIG. 6.6. The flux density at the image on the retina depends on the angle V subtended by the radius of the exit pupil of the eye, E_e' .

being the angle of projection in the eye, as shown in Fig. 6.6. The formula shows that the subjective brightness is independent of the distance of the object in accordance with experience, the angle V being relatively constant.

If an optical instrument is used in front of the eye, the subjective brightness is affected by a possible reduction in V if the exit pupil of the instrument is smaller than the entrance pupil of the eye. The eye is usually placed at the exit pupil E' of the instrument for the purpose of obtaining the maximum field of view. Let ρ' be the radius of this exit pupil, and let h be the radius of the pupil of the eye. If $d = -p'$ is the distance of the final image in the instrument, $d > 25$ cm, it is apparent, Fig. 6.7, that

$$\rho' = d \tan U' \doteq d \sin U'$$

since the angle U' is small.

Two cases must be considered. (1) If $\rho' > h$, the latter becomes the exit pupil of the instrument and eye combination, and the pupil is filled with light. In this case, the image illuminance on the retina is unchanged and one obtains **normal brightness**. However, (2) if the exit pupil is smaller than the eye pupil, $\rho' < h$,

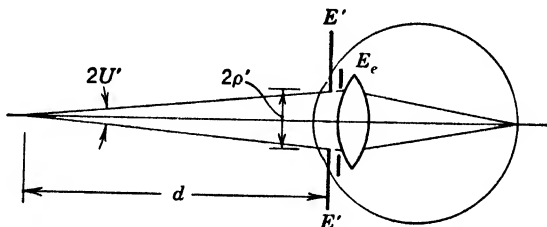


FIG. 6.7. Relation between the size of the exit pupil E' of an instrument, the angle of projection $2U'$, and the image distance d .

the illuminance H_0 is reduced to a smaller illuminance H' because of the reduced angle of projection in the eye, V' , so that

$$\frac{H'}{H_0} = \frac{\sin^2 V'}{\sin^2 V} = \frac{\text{area of exit pupil}}{\text{area of eye pupil}} = \left(\frac{\rho'}{h}\right)^2$$

This formula holds only if $\rho' \leq h$ and $H'/H_0 \leq 1$, for H'/H_0 can in no case be greater than unity as pointed out under case (1).

6.5 Normal Magnification

The above formula can be applied to investigate the conditions under which a microscope may give normal brightness. The maximum magnification giving $H' = H_0$ is called the **normal magnification**. Since $\rho' = d \sin U'$, it is found that

$$\frac{H'}{H_0} = \frac{d^2 \sin^2 U'}{h^2} = \frac{d^2 (n \sin U)^2}{h^2 n'^2 \beta^2}$$

Thus, for any given numerical aperture, there is a maximum magnification for which there is no reduction in brightness. To find its value, set $H'/H_0 = 1$ and let $n' = 1$ for reasons given previously. The normal magnification is then

$$\beta_n = \frac{d}{h} (n \sin U)$$

Now d is the distance of the final virtual image and should be taken equal to 250 mm, the normal reading distance, while h is on the average equal to 2 mm. Hence

$$\beta_n \doteq 125 (n \sin U)$$

Since a numerical aperture of 1.5 can scarcely be exceeded, we have an upper limit of about 200 diameters for the normal magnification. No microscope can magnify more than that without a reduction in image brightness which is thereafter proportional to the inverse square of the magnification. The importance of a

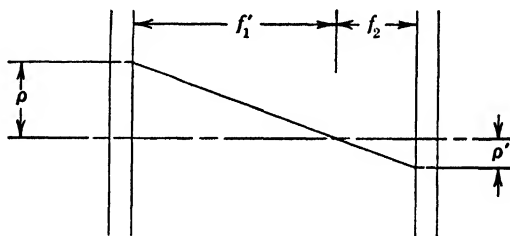


FIG. 6.8. Determination of the radius ρ' of the exit pupil of a telescopic system.

large numerical aperture in obtaining bright images is also apparent from the formula.

In telescopic systems, similar conditions apply, but they are expressed more simply in terms of linear apertures. In Fig. 6.8 ρ and ρ' represent the radii of the entrance and exit pupils, and f_1' and f_2 are the focal lengths of the objective and eyepiece respectively. By similar triangles it is seen that

$$\frac{\rho}{\rho'} = \frac{f_1'}{f_2}$$

It was proved in Art. 3.7 that this ratio of the focal lengths is equal to the angular magnification. Therefore

$$\rho = \gamma \rho'$$

In the expression $H'/H_0 = \rho'^2/h^2$, the exit-pupil radius may be eliminated by the preceding formula, showing that

$$\frac{H'}{H_0} = \frac{\rho^2}{\gamma^2 h^2}$$

If $\rho > \gamma h$, the full aperture of the telescope is not utilized, since the eye pupil becomes the exit pupil of the entire system. One then obtains normal brightness in an ideal telescope having no internal light losses. However, when $\rho < \gamma h$, the brightness is less than normal, in accordance with the above formula. This formula shows that one may compensate for higher magnifying power γ by a corresponding increase in the radius ρ of the entrance pupil. The most efficient combination of eye and telescope, which still gives normal brightness, is one in which the exit pupil of the telescope has the same radius as the pupil of the eye. Since $H'/H_0 = 1$ for this case, the normal magnification is found to be

$$\gamma_n = \frac{\rho}{h}$$

The radius of the pupil of the eye varies from about 1 mm in bright light to about 3.5 mm at low intensities. Thus the diameter of objective giving normal brightness of image is

$$2\rho = 2h\gamma = 2\gamma \text{ mm in bright light}$$

and

$$2\rho = 7\gamma \text{ mm in dim light, as at night}$$

Therefore a 7-power telescope, for example, should have an objective of at least 14 mm for daytime use, and at least 49 mm for night-time use. The popular 7 x 50 binocular fulfills both requirements. However, if used exclusively for daytime observations, the aperture may just as well be smaller, and the instrument less expensive, without impairing its optical performance.

One might summarize the requirements for normal magnification by stating that, as the area of the retinal image is increased as γ^2 , the area of the entrance pupil must be correspondingly increased as γ^2 to keep the flux density constant. However, if the object is a point source, its retinal image will be a diffraction disk whose area is independent of γ . In this case, an increase in entrance pupil will increase the brightness above the normal value so that $H'/H_0 = \rho^2/h^2$ may exceed unity. It is for this reason that stars can readily be seen in daytime if a large enough telescope is used. The star brightness is then greater than with the unaided eye, while that of the background is at best the same as with the unaided eye. Thus it becomes clear that the new 200-in. telescope

will enable fainter celestial objects to be seen or photographed than is possible with the 100 in. using a photographic emulsion which is four times as fast.

PROBLEMS

1. What is the illuminance on a workbench midway between two lamps of 150 cp 4 ft above the bench and 6 ft apart, assuming that twice as much light comes to the surface indirectly from the walls and ceiling as directly from the lamps?

2. The parabolic reflector of a searchlight reflects 50 per cent of the light flux from a 5000-cp lamp placed at its focus. What is the illuminance produced by the searchlight at 500 ft if the reflected beam illuminates an area 20 ft in diameter at this distance?

3. A 200-cp lamp illuminates a photoelectric cell of 2 sq in. area at a distance of 36 in. (a) What is the illuminance at the cell in foot-candles? (b) What is the light flux to the cell in lumens?

4. What is the illuminance on the ground half-way between two lamps of 2000 cp each, the distance between lamps being 30 ft and their height above the ground 15 ft?

5. What is the luminous emittance of a diffusely reflecting surface whose reflectance is 0.65 and illuminance is 25 foot-candles? (a) Give the result in millilamberts. A surface has a luminous emittance of 1 **lambert** if it reflects diffusely 1 lumen of light flux per square centimeter. (b) Give the luminance in candle power per square centimeter.

6. With a lamp 20 cm from an object, an exposure of 1 sec is needed to photograph it, using a relative aperture of $f/4$. What time of exposure is needed if the lamp distance is changed to 40 cm and the aperture is changed to $f/16$?

7. A camera having an $f/4$ lens is used to copy black-and-white line drawings at a distance of 4 ft with a magnification of $1/3$. What is the maximum illuminance on the film in foot-candles if the luminous emittance of the white areas is 100 millilamberts and there is a 20 per cent light loss in the lens?

8. A motion-picture projector contains a high-intensity arc of 85,000 candles per sq cm and projects a 9 x 12-ft picture at a distance of 80 ft. The exit pupil of the projection lens is 2 in. in diameter. What is the maximum illuminance at the center of the screen if the optical system is 60 per cent efficient?

9. An absorbing screen 4 cm thick transmits 50 per cent of a certain radiation. (a) What percentage would be transmitted by a thickness of 2 mm, neglecting surface reflection? (b) What would be the answer if there were a reflection loss of 4 per cent at each surface?

Chapter 7. Optical Instruments

7.1 The Eye

The eye is not solely an optical instrument, for it is nourished and controlled by a living organism, and the images in it are converted into nerve impulses travelling along the optic nerve to the brain, where their interpretation involves a great deal of psychology. Here we will concern ourselves with only that portion of the subject that is necessary for the understanding of optical aids to vision. For other aspects of vision, reference is made to other sources, for example, Walsh's *Photometry* and Mueller-Pouillet's *Lehrbuch der Physik*. The cardinal points of the eye as an optical system are shown in Fig. 7.1. The center of the transparent window of the eye, the **cornea**, is at V , and the principal foci with relaxed accommodation are at F_0 and F_0' . Of the two focal lengths, f and f' , the latter is the greater because of the greater index of refraction of the medium on the image side of the

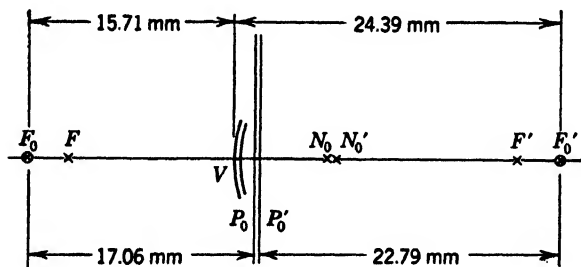


FIG. 7.1. Cardinal points and focal lengths of a normal human eye. The anterior vertex of the cornea is at V .

optical system. The nodal points are at N_0' and N_0 and are located at a distance f from F_0' and at f' from F_0 respectively. The optical system consists essentially of the cornea and a so-called **crystalline lens**, the front surface of which is 3.1 mm behind the front surface of the cornea. The greater part of the focussing of light rays takes place at the cornea because of the relatively large

change in index of refraction combined with the small radius of curvature. The crystalline lens is not entirely homogeneous but has an average index of refraction of 1.42 and a thickness of 3.6 mm. The index of refraction of the aqueous humor between it and the cornea is 1.337, and the index of the vitreous humor behind it is also about 1.336, the power of the lens being +19.11 diopters when accommodation is relaxed. The power of the cornea is about 43.05 diopters, that of the cornea and lens together being 58.64 diopters. In a normal or **emmetropic** eye, the retina is at the second principal focus F_0' , so that distant objects are in focus on it when there is no accommodation.

Accommodation takes place through the contraction of a ring of muscle around the lens which causes its power to increase. A study of the images reflected by the various surfaces shows that the front surface of the lens increases in curvature much more than the back surface, and that the cornea remains unchanged. For a normal or average eye, the process of accommodation shifts the principal foci to the points F and F' , Fig. 7.1, without changing the locations of the principal planes and nodal points by more than half a millimeter. In the state of maximum accommodation, images of objects at about 25 cm are projected on the retina.

The retina is covered with nerve endings of two distinct types, called **rods** and **cones** after their appearance. There are about 120,000,000 nerve endings, and they are most closely packed at the **fovea**, which is a small spot slightly off the optical axis of the lens system. There are no rods, only cones, at the fovea. The detail in that part of the image projected on the fovea can be best resolved, the maximum resolution corresponding to a visual angle of slightly less than one minute of arc in bright light. The resolution falls off as the intensity of the light is diminished because of an effective increase in the coarseness of the retina, which is due to the fact that some of the cones require a relatively high threshold energy for stimulation. The fovea measures only 0.25 mm across and subtends an image angle of 0.70° , which corresponds to an object angle of about 0.87° . Around it is a region called the **macula** of about 7° in extent, in which the resolution is only slightly inferior to that in the fovea. The rest of the retina is of increasing coarseness and serves more for the detection of the presence of objects than for their close examination.

The **amplitude of accommodation** may be represented by the interval between the object point in focus when the accommodation is relaxed, called the **far point**, and the object point in focus when the maximum accommodation is exerted, called the **near point**. For an emmetropic, or normal, eye, these points are respectively at infinity and at 25 cm from the eye. The locations of the far point and of the near point vary with age as the crystalline lens loses its elasticity. The distance of the near point increases steadily with age from about 10 cm at birth to about 33 cm at an age of 80 years. The far point remains practically unchanged until an age of 50 years and then gradually diminishes to about 33 cm at 80 years, after which there is practically no range of accommodation. This change in far point and loss of accommodation is called **presbyopia**.

7.2 Spectacle Lenses

The widespread use of spectacle lenses shows that eyes frequently suffer from defective accommodation, a condition known as **ametropia**. The principal forms of ametropia are **myopia**,

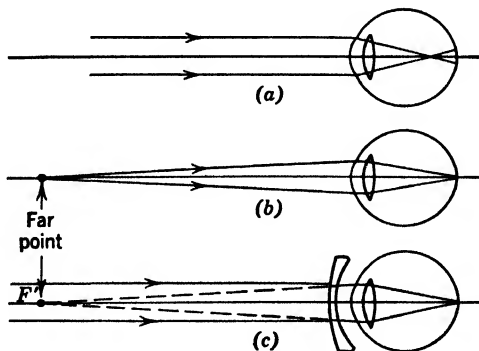


FIG. 7.2. Myopia. (a) Incident parallel rays focus in front of the retina. (b) With zero accommodation, rays from the far point come to a focus on the retina. (c) A diverging lens whose principal focus F' is at the far point provides ray pencils which focus on the retina with zero accommodation, the object being at infinity.

or nearsightedness, **hyperopia**, or farsightedness, **astigmatism**, and perhaps **heterophoria**, or muscular imbalance.

Myopia is often due to an abnormally long eyeball which places the retina beyond the principal focus, Fig. 7.2(a). In this case,

objects must be moved closer to the eye until their image falls on the retina. When accommodation is relaxed, this point is called the far point, and when maximum accommodation is exerted, it is called the near point. These are therefore the optical conjugates of points on the retina in the respective cases. A consideration of Fig. 7.2 shows that in order to focus the parallel rays from a distant object when the accommodation is relaxed, a negative lens having its principal focus at the far point must be used. The back surface of this lens is usually placed at 13 to 15 mm from the vertex of the cornea in order to clear the lashes and to keep the magnification unchanged, as will be explained later. Thus the proper vertex focal length is equal to the distance from the cornea to the far point minus the distance from the cornea to the rear vertex of the spectacle lens. In elementary discussions, the latter distance is often neglected, and so is the distinction between vertex power and optical power, but these refinements become important if the lens is strong.

As an example of myopia, suppose that the far point is at a distance of 200 mm from the cornea. The vertex focal length of the spectacle lens required is then -185 mm, the vertex power being -5.4 diopters. The near point with the spectacle lens may be found by applying the thin-lens formula

$$\frac{1}{s} = \frac{1}{f} - \frac{1}{s'}$$

where f is the focal length of the spectacle lens, and s' is the image distance. In this case s' is negative, since the spectacle-lens image is virtual, and the magnitude of s' is equal to the distance from the second principal plane of the lens to the near point of the unaided eye. Let this be 100 mm, i.e., about 115 mm from the cornea. Then, assuming that the optical focal length is 5 mm greater than the vertex focal length,

$$\frac{1}{s} = \frac{1}{-190} - \frac{1}{-100} = \frac{9}{1900}$$

from which one finds that $s' = 211$ mm or 21.1 cm. With the object at this distance from the first principal plane of the lens, the rays enter the eye coming from an image at the location for which the eye can focus by accommodation.

It may be more convenient to use the vertex focal length in the spectacle-lens formula instead of the optical focal length, but one must then measure the distance s' from the rear vertex, and s from the conjugate of the rear vertex, which is approximately one third the thickness of the lens toward the object.

Compact, direct-reading instruments have been devised for determining vertex diopters over a wide range of positive and

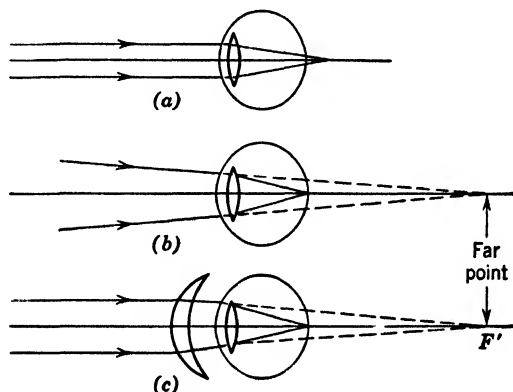


FIG. 7.3. Hyperopia. Zero accommodation in all three diagrams. (a) Incident parallel rays are intercepted by the retina before coming to a focus. (b) Ray pencils converging toward the far point come to a focus on the retina. (c) A converging lens with its principal focus F' at the far point provides ray pencils which focus on the retina with zero accommodation, the object being at infinity.

negative powers. These employ an illuminated target, a strong positive auxiliary lens, and an eyepiece. They are called **Diop-trometers, Vertometers, Focometers, or Lensometers**, depending on the manufacturer. The optical system of one standard form of instrument is described in Hardy and Perrin's *The Principles of Optics*, page 371.

A hyperopic eye generally has a flattened eyeball so that the retina lies forward of F_0' . Thus, when relaxed, it can focus on the retina rays converging toward some point behind the eye, as shown in Fig. 7.3(b). To focus even parallel rays requires some accommodation. As before, to focus parallel rays without accommodation, a lens must be used whose principal focus lies at the far point. Consequently, its power is the reciprocal of the dis-

tance of the far point behind the eye, increased by the distance from the eye to the lens. The lens is obviously converging or positive.

The new near point of the eye with the lens in place may be found by applying the lens equation to the spectacle lens, as before, using the distance to the near point of the unaided eye as a virtual (negative) image distance.

Astigmatism is generally due to a non-spherical cornea, although sometimes the crystalline lens is astigmatic. If the ciliary muscles do not act equally around the periphery of the lens, the amount of astigmatism will change with accommodation. This effect is often overlooked. Astigmatism causes the image of a point to consist of two lines at right angles to each other, as shown in Fig. 4.6, although if the lens is aspherical, the effect will be observed even when the object points are on the principal axis. The presence of astigmatism is detected by the use of a cross or a star-shaped object, the images being out of focus unless an arm of the object is in the same direction as one or the other of the two line images of a point. The lens system has a maximum power in a direction at right angles to the line focus which is closest to the lens. To correct for astigmatism, a non-spherical lens having appropriately different powers in the two principal meridians is required.

The far-point (and similar near-point) diagrams of Figs. 7.2 and 7.3 can be generalized to include astigmatism if one introduces two far points (and two near points) corresponding to the principal meridians of maximum and minimum power. The required lens powers can then be found for each meridian. For example, if the far point for lines in the vertical meridian (axis 90°) is at 2 meters and for lines in the horizontal meridian (axis 0°) is at 1 meter, the required lens must have a power of -0.5 in the 0° meridian and -1.0 diopter in the 90° meridian. Note that the power of a cylinder acts in a meridian at right angles to its axis. Hence the required powers may be obtained by a sphero-cylindrical combination of either -0.5 D Sph $- 0.5$ D Cyl axis 0° or by -1.0 D Sph $+ 0.5$ D Cyl axis 90° . The change from minus to plus cylinders or vice versa in a prescription is called **transposition**.

It is sometimes necessary to combine a prism with an ophthalmic lens. If the prism is not too strong, it may be obtained by de-centering the lens. If the center of the lens is displaced by a distance x from its optical center, which is on a line joining the two

centers of curvature, the angle of deviation is seen to be $\tan^{-1} xD$, Fig. 7.4. Since the prism strength in **prism diopters** Δ is measured by the centimeters of displacement of the transmitted ray at a distance of 1 meter, we have, by similar triangles,

$$\frac{\Delta}{100} = \frac{x}{1/D} \quad \text{or} \quad \Delta = 100xD$$

x being in meters, or simply $\Delta = xD$ if x is in centimeters. This formula is known as Prentice's rule. An image seen through a prism is displaced in a direction away from the base. Thus a

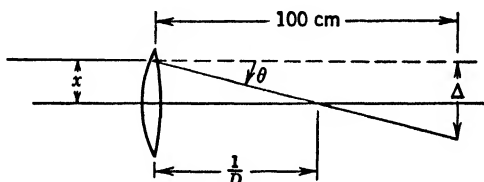


FIG. 7.4. Relation between the amount of prism Δ and the amount of decentering x of the lens. Prentice's rule.

base-down prism causes an image to be seen higher than without the prism. Obviously, decentering a negative lens induces a prism with the base in the opposite direction to that induced by the same decentering of a positive lens.

As the eye rotates behind a spectacle lens, the utilized rays pass more and more obliquely through the lens. This introduces an increasing amount of astigmatism, which impairs vision and limits the useful field of view. In the case of a symmetrical biconvex or biconcave lens, the useful field extends to only 8° from the principal axis. A plano-convex or plano-concave lens with the flat side toward the eye has a useful field of about $\pm 12^\circ$, and this is further increased by the use of meniscus lenses.

The best shape of meniscus lens depends on its power in the manner shown graphically by the **Tscherning ellipse** in Fig. 7.5. The abscissas are the powers of the lens, and the ordinates are the powers of the front surface $D_1 = (n - 1)/R_1$. In general, there are two shapes of meniscus lens that may be used, the shallow meniscus defined by the lower arc or the **Ostwald branch** of the ellipse, and the deep meniscus lens given by the upper or **Wollaston branch**. The Ostwald branch is more commonly used

nowadays. As the ellipse shows, only lenses having powers between -22.5 and $+7.5$ diopters can be completely corrected for astigmatism. Optical shops usually stock lens blanks with one surface finished. In order to surmount the impossibility of stocking an infinite variety of finished surfaces, the lower arc of the ellipse is approximated by six or more short horizontal lines for the various powers. A completely corrected lens is obtained only for powers corresponding to the intersections of the lines with the el-

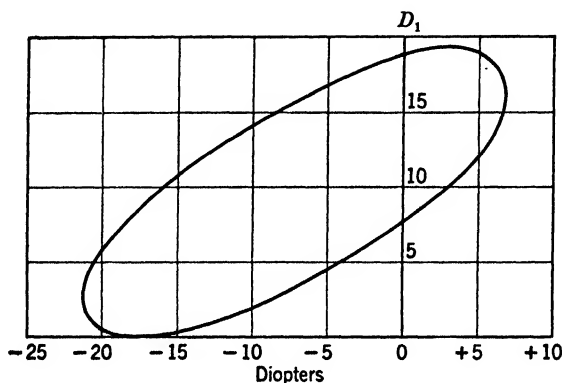


FIG. 7.5. Tscherning ellipse. Front surface powers D_1 of spectacle lenses of total power D when corrected for astigmatism.

lipse. Weaker or stronger lenses will show a small amount of astigmatism, but as this becomes too large, the next standard curvature is used.

The magnification of the image on the retina is given by the formula

$$M = -\frac{f}{x} = -\frac{1}{xD}$$

where D is the combined power of the eye and the spectacle lens. According to the standard formula for combinations of two optical systems, $D = D_E + D_S - dD_ED_S$, where D_E is the power of the eye and D_S is that of the spectacle lens. The ratio of the magnification with the lens to that of the unaided eye, assuming that x is so large that we may neglect variations in its value due to the displacement of the principal focus, is

$$\frac{M}{M_E} = \frac{D_E}{D_E + D_S(1 - dD_E)} = \frac{1}{1 + D_S(f_E - d)} = \frac{1}{1 - pD_S}$$

where p is the distance of the second principal plane of the spectacle lens in front of the first principal focus of the eye, Fig. 7.6. It is readily seen that, if $p = 0$, it follows that $M/M_E = 1$, and the size of the image with any power D_s is the same as for the unaided eye. In general, this defines the correct location of the spectacle lens. If the lenses for both eyes are of the same power, there is no harm in mounting the lens somewhat farther or nearer to the eye. If, however, this is not the case (a condition known as **anisometropia**) any shift in the distance of the lenses will induce a condition of unequal image size known as **aniseikonia**. Sometimes this is present without spectacle lenses. The remedy is then to use lenses whose principal planes P' are at appropriately different distances from the eye. Instead of changing the spectacle lens location, one may change its thickness or depth of curve to accomplish the same placement of the second principal plane of the lens. The presence of astigmatism complicates the problem by requiring a simultaneous solution of the problem for the two principal meridians. Aniseikonia is not always due to an actual difference in the sizes of the retinal images, but may be due to unequal densities of the rods and cones or of the perceptual centers in the brain. Thus images may appear unequal in size without actually being so. This defect may also be compensated by an appropriate displacement of P' .

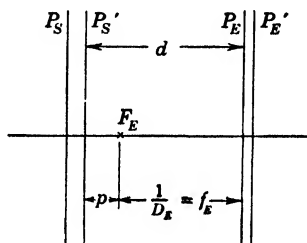


FIG. 7.6. Magnification by spectacle lenses depends on the distance p .

7.3 The Simple Magnifier

The size of image on the retina is proportional to the visual angle α subtended by the object O at a distance d from the eye. Because of limited accommodation, the angle α will have a maximum value when O is at the near point of the eye. The finest structural detail of the image can be resolved under this condition. If O is assumed to be a small object or a small portion of a large object, Fig. 7.7, the visual angle is given by $\alpha = O/d$. With the aid of a positive lens, a virtual image can be created within the range of accommodation of the eye when the object is slightly within the principal focus of the lens. The image subtends a larger

angle at the lens, $\beta = O/s$. The eye is assumed to be close to the lens, so that β is practically the same as the visual angle for the

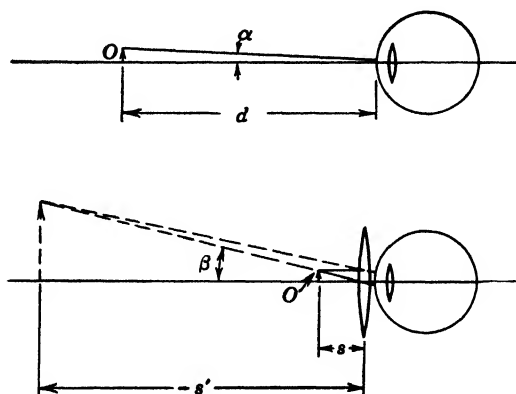


FIG. 7.7. Magnifying power of a simple magnifier or simple microscope.

enlarged image. The magnification ratio of the two retinal images is approximately equal to the ratio of the visual angles, or

$$M = \frac{\beta}{\alpha} \doteq \frac{\tan \beta}{\tan \alpha} = \frac{O/s}{O/d} = \frac{d}{s}$$

The distance d is usually taken to be 25 cm, the normal near-point distance for the emmetropic eye. The value of $1/s$ is given by the lens formula, so that

$$M = 25 \left(\frac{1}{f} - \frac{1}{s'} \right) = \frac{25}{f} - \frac{25}{s'}$$

The numerical value of s' lies between 25 cm and infinity. Taking the former value, one finds, since s' is negative, that

$$M = \frac{25}{f} + 1$$

If the virtual image is at infinity, one has simply

$$M = \frac{25}{f}$$

For the larger powers or shorter focal lengths there is little practical difference between these two results, so that the second is often

taken as giving the approximate magnification over the entire range of image locations. Moreover, the latter corresponds to the location of the image in the absence of accommodation, which is the condition of most comfortable vision. Note that M is quite different from s'/s , which increases to infinity as s approaches f , while M decreases by unity under the same condition. Although one may speak of magnification in terms of "diameters," visual magnification is fundamentally a ratio of visual angles. This angular ratio becomes equal to the transverse linear magnification only if the value of s' is 25 cm, which is rarely the case.

The magnifier theory differs from that of spectacle-lens magnification, Art. 7.2, in that there is a great change in the object distance x in the present case instead of a negligible one. In the discussion of spectacle lenses, the sizes of images were compared for a given object location whether or not they were in focus on the retina. In the case of strong lenses, the magnifier theory must be applied, since the object distance must be materially reduced when the lens is used.

7.4 Eyepieces

Eyepieces are fundamentally magnifiers designed to provide a more perfect image than can be obtained by means of a single lens of the same power. The discussion of aberrations of lenses and their reduction is applicable here. In this connection it should be noted that an eyepiece must have a large angular field in order that the field of the objective is not unduly restricted. In any case this will be $1/M$ times the field of the eyepiece, where M is the overall magnifying power of the instrument.

The most common eyepiece on microscopes is the **Huygens eyepiece**, a cross-section of which is shown in Fig. 7.8(a). The first lens, or **field lens**, usually has a focal length three times as great as that of the **eye lens**. The separation is twice the focal length of the eye lens to satisfy the condition for partial lateral achromatization, Art. 4.9,

$$d = \frac{f_1 + f_2}{2}$$

The lens curvatures and glass types are computed to minimize spherical aberration, coma, and astigmatism. This type of eyepiece requires that the incoming rays be convergent, forming a

virtual object, in order that the final image be in the range of accommodation of a normal eye. Hence it cannot be used as a magnifier of real objects. If a scale or grid is to be used for measuring or counting, it must be placed at the location of the first real image between the field and eye lenses. This leads to some inaccuracy due to relative distortion, because the scale is magnified by only the eye lens, whereas the image projected on it is magnified by the entire lens system, the distortions in the two cases usually not being the same.

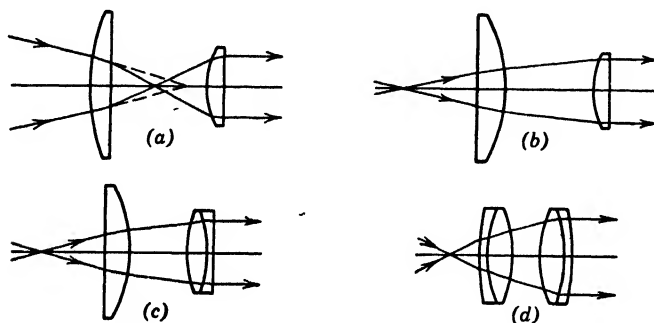


FIG. 7.8. Types of oculars: (a) Huygens eyepiece, (b) Ramsden eyepiece, (c) Kellner eyepiece, (d) orthoscopic eyepiece.

It is for this reason that measuring instruments generally employ a positive eyepiece of the Ramsden, Kellner, or similar type. The **Ramsden eyepiece**, Fig. 7.8(b), also employs two lenses, but these are often equal in focal length and separated by three fourths of the focal length of either. The separation thus differs from the optimum, which is half the sum of the focal lengths. Otherwise, the images of the surfaces of the field lens would be superimposed over the image of a distant object, and any dust or imperfection of surface would be very disturbing. Hence the principal focus is displaced to a point conveniently located in front of the field lens. The field stop, reticle, or cross-hairs are located here.

The **Kellner eyepiece**, Fig. 7.8(c), offers a considerable increase in image quality and is widely used in higher-power telescopes, gun sights, and in fire-control instruments. It also serves as an excellent magnifier for examining spectrograms and fingerprints and for other general use. The symmetrical doublet, Fig. 7.8(d), is another example of a high-quality eyepiece or magnifier.

7.5 The Compound Microscope

For obtaining magnifying powers of more than thirty or so diameters with a satisfactory quality of image, a more complex lens system is required. The compound microscope consists of two distinct lens systems: the **objective**, which creates a magnified real image of the object, and the **eyepiece**, which further magnifies this real image, Fig. 7.9. The overall magnifying power is the product of the magnifying powers of the objective and eyepiece. The former, disregarding its sign, is given by the usual formula

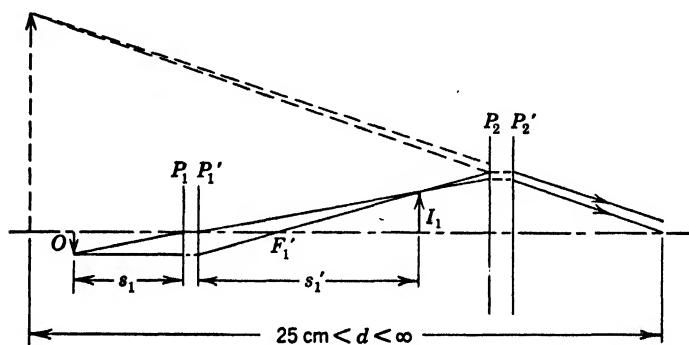


FIG. 7.9. Optical system of a compound microscope.

$$M_0 = \frac{s_1'}{s_1} = \frac{s_1'}{f_0} - 1$$

If the focal length f_0 of the objective is quite small compared with the image distance or optical tube length L , one may write the approximation that

$$M_0 = \frac{L}{f_0}$$

The overall magnifying power is then

$$M = M_0 M_e = \frac{25L}{f_0 f_e}$$

It is evident that a high magnifying power may be obtained by the use of lenses of short focal length, of which the objective is generally the shorter by perhaps a factor of ten or so. The length L is usually not over 16 cm for the sake of a convenient compactness

of the instrument. A high-power objective may have a focal length as small as 2 mm. The cross-sections of a few typical objectives are shown in Fig. 7.10.

The brightness of the image and the resolving power of the instrument are dependent on the numerical aperture, this being the product of the index of the immersing medium and the sine of half the angular aperture, so that (see Fig. 7.10)

$$\text{N.A.} = n \sin A$$

The smallest separation of points or lines that can be resolved is limited by the diffraction of light, Art. 11.4, and amounts to

$$d_{\min} = \frac{1.22\lambda}{2 \text{ N.A.}}$$

It is principally because of the short wave length of light, about 0.5 micron, that a good microscope allows one to see fine structural detail. In order to utilize to the utmost the resolving power ob-

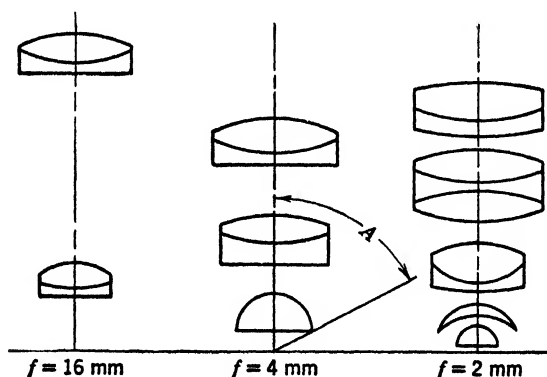


FIG. 7.10. Types of microscope objectives. The angle A defines half the angular aperture.

tainable with a given wave length, the numerical aperture must be as high as possible. This requires a medium of high index of refraction above and below the object and also a large angular aperture. The index of the immersing medium is, however, limited to that of the glass of which the front component of the objective is made. Because of this and the inability to use an angular aperture of quite 180° , the numerical aperture is never much more than 1.5. Thus the finest detail resolvable is limited

to about one third of the wave length of the light used. With microscopes for visual observation, this is about 0.18μ . Microscopes employing ultraviolet light and photographic registration may have a resolving power of nearly twice that of the visual. The reader interested in practical microscopy is referred to Shillaber's *Photomicrography*. Incidentally, electron microscopes achieve their high resolving power because of the short effective wave length of an electron beam (about 5×10^{-10} cm) in spite of their low numerical aperture of about 0.02.

Whatever type of microscope or any other instrument one employs, its resolving power indicates merely the finest detail reproduced in the image. Since it is the image which is examined, not the object, the finest detail contained in the image is the finest detail one can see, however much one may magnify the image. However, the magnifying power is also important because the eye can resolve only structures which subtend at least one minute of arc at the eye, or about 0.073 mm at a distance of 25 cm. Hence whatever detail is to be perceived must be magnified to at least this value. Thus the finest detail resolvable by an optical microscope, about 0.18μ , must be magnified at least 400 times to be perceived directly. A somewhat higher magnification is often convenient, but it cannot bring out any new detail. Higher magnifications are said to be "empty." On the other hand, one may profitably use magnifications of 100,000 with an electron microscope if the "lenses" are properly designed.

7.6 Telescopes

A real image of a distant object is formed by the objective lens at a point which is slightly beyond its principal focus. This real image is magnified by one of the eyepieces described previously. The objective is corrected for chromatic aberration, spherical aberration, and coma. It is not necessary to consider very oblique rays, since the angular field is necessarily small, being $1/M$ times that of the eyepiece, M being the magnification. Thus, if the eyepiece has a useful field of 60° and the magnification is 15, the field of the telescope objective will be 4° . Figure 7.11 shows the cardinal points of the components of an astronomical telescope. The magnifying power is

$$M = \frac{\beta}{\alpha}$$

If we consider a small object, or some small detail of a larger object, the corresponding angles α and β may be represented approxi-

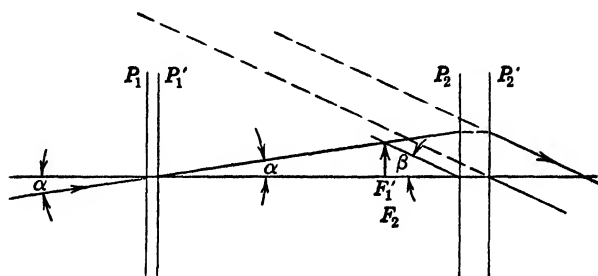


FIG. 7.11. Optical system of an astronomical telescope. The arrow at F_1' indicates the size of the first image I_1 .

mately by I_1/f_1 and I_1/f_2 respectively. Thus

$$M = \frac{I_1/f_2}{I_1/f_1} = \frac{f_1}{f_2}$$

Since the magnifying power of the eyepiece is $25/f_2$, the contribution of the objective is $f_1/25$. It is apparent that, if one is to divide a high overall magnifying power between the objective and the eyepiece, the focal length of the objective must be quite large. For hand-held telescopes, this need not be the case, since the

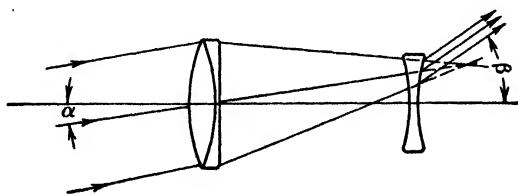


FIG. 7.12. Galilean telescope.

highest power that one may use is about eight, and most of this can be provided by the eyepiece.

A telescope with only two positive lens systems, as shown in Fig. 7.11, gives inverted images, since the rays enter the eye from the opposite direction with the telescope than without it. If a diverging lens is used as an eyepiece, the image will be rectified, as shown in Fig. 7.12. This type is called a **Galilean telescope** and is in

common use in opera glasses and in lower-priced field glasses. A disadvantage of this system is that the field of view is strongly limited at the higher powers. Moreover, it is not so easy to design a good-quality diverging eyepiece as a positive one. Hence the usefulness of this optical system is restricted to the lower powers.

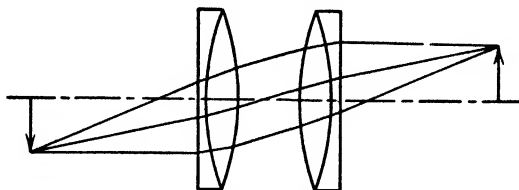


FIG. 7.13. Erecting lens system.

If there is no objection to increasing the length of the telescope, a set of erecting lenses, as shown in Fig. 7.13, may be used between the objective and eyepiece. The trench periscope, for example, is equipped with this type of erecting system. It is also commonly used in semiportable or fixed terrestrial telescopes which are mounted on stands.

For maximum compactness a set of totally reflecting prisms is employed. The operation of such a pair of **Porro prisms** is illustrated in Fig. 7.14. Figure 7.14(a) shows that the rays which would form the upper part of the image are caused to form the lower part. Figure 7.14(b) shows that no such rectifying effect takes place in the perpendicular direction. Since the image is not only inverted but also reversed, as is evident from the axis of symmetry of the objective lens, a second prism must be used at right angles to the first to obtain complete rectification. It will be noted that the use of prisms leads to a more compact form of instrument and that it naturally lends itself to an increase in the distance between the entrance pupils of a binocular instrument, thus enhancing the stereoscopic effect. Various other forms of telescopes using prisms are described by Jacobs in *Fundamentals of Optical Engineering*.

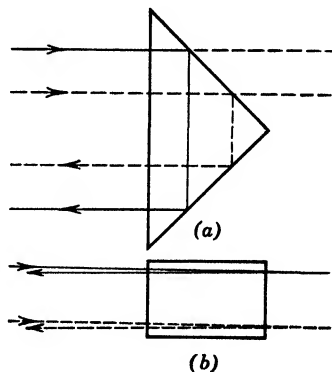


FIG. 7.14. Porro prism.

7.7 Reflecting Telescopes

The largest telescopes employ parabolic mirrors as objectives. This construction has the advantage that only one surface need be carefully ground and polished, and that the glass blank need not be free of striae, ripples, and other flaws. The image produced by the parabolic objective is usually reflected to one side by a hyperbolic mirror placed near the focus, or it may be reflected back through a hole in the center of the parabolic objective. Spherical

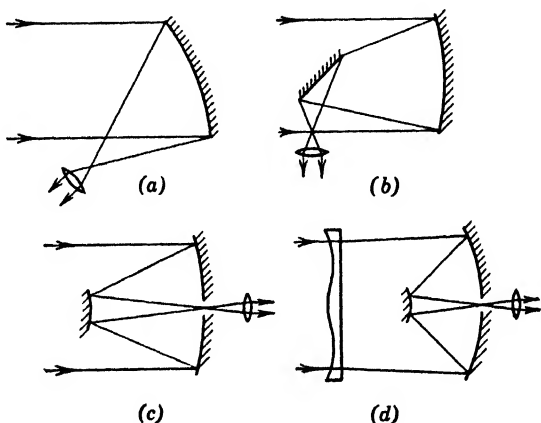


FIG. 7.15. Reflecting telescopes: (a) Herschelian, (b) Newtonian, (c) Cassegrainian, (d) Schmidt type.

mirrors are generally not used because of their large spherical aberration and coma. An exception is the Schmidt type of telescope, which employs a spherical mirror and a compensating plate in the form of an irregularly shaped diverging lens at the center of curvature of the reflector. This type of telescope is often made with a large relative aperture. The principal types of reflecting telescope, diagrammed in Fig. 7.15, are the (a) **Herschelian**, (b) **Newtonian**, (c) **Cassegrainian**, and (d) **Schmidt type**. The chief advantage of the reflecting telescope, besides its relative ease of construction, is that it has no chromatic and spherical aberration and can therefore be made with a large relative aperture of $f/5$ or so. Schmidt telescopes of up to $f/0.9$ have been constructed.

7.8 Prism Spectroscopes and Spectrographs

An aperture and prism were first used by Newton in 1666 to analyze beams of light into their constituent colors or wave lengths. The resolving power was low, so that spectrum lines were not discovered until early in the nineteenth century by Wollaston with a slit, prism, and lens combination. In 1819 Fraunhofer used a narrow slit, prism, and telescope in measuring the indices of refraction

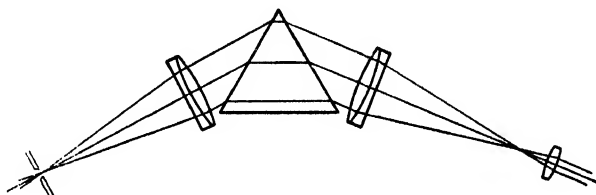


FIG. 7.16. Prism spectroscope.

of glass for his lenses. With this arrangement he discovered many sharp lines in emission spectra and in absorption spectra. The addition of a collimating lens between the slit and the prism seems to be due to Babinet in 1839. This optical system, which is shown in Fig. 7.16, is still widely used in present-day spectroscopes and spectrometers.

When the telescope is replaced by a camera, the instrument is called a **spectrograph**. Sometimes the prism is replaced by a diffraction grating. The theory of the grating, an invention of Fraunhofer's, will be discussed in Art. 11.6. The theory of the prism follows.

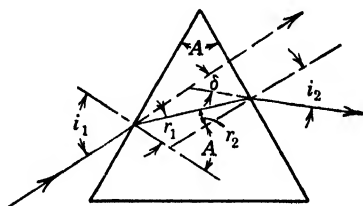


FIG. 7.17. Deviation of a light ray by a prism.

Figure 7.17 shows a light ray refracted at the two surfaces of a prism, the prism angle being A . The light is deviated through an angle δ , where

$$\delta = (i_1 - r_1) + (i_2 - r_2)$$

Since the angle A of the prism is equal to the angle between the normals to the surfaces

$$A = r_1 + r_2$$

by the theorem that the exterior angle of a triangle equals the sum of the two opposite interior angles. It also follows that

$$\delta = i_1 + i_2 - A$$

The angles i and r are related by Snell's law

$$\sin i_1 = n \sin r_1 \quad \text{and} \quad \sin i_2 = n \sin r_2$$

where n is the index of refraction of the prism with respect to the external medium, usually air.

These equations may be used to trace any ray through a prism. Two special cases are of interest. If A and i_1 are small, one may write

$$i_1 = nr_1 \quad \text{and} \quad i_2 = nr_2$$

so that

$$\delta = (n - 1)A$$

This result is quite independent of i_1 as long as this angle is not too large, making this a useful formula for the deviation by weak prisms.

For larger angles one finds that, as i_1 is varied, δ passes through a minimum. The condition is readily found by the usual method, for then

$$\frac{d\delta}{di_1} = 0 = 1 + \frac{di_2}{di_1}$$

Because $i_2 = \sin^{-1}(n \sin r_2)$, it follows that

$$\frac{di_2}{di_1} = \frac{n \cos r_2}{\sqrt{1 - n^2 \sin^2 r_2}} \frac{dr_2}{di_1}$$

Since $\sin i_1 = n \sin (A - r_2)$, and therefore

$$di_1 \cos i_1 = -n \cos (A - r_2) dr_2$$

one finds that

$$\frac{di_2}{di_1} = - \frac{\cos r_2}{\sqrt{1 - n^2 \sin^2 r_2}} \frac{\cos i_1}{\sqrt{1 - \frac{\sin^2 i_1}{n^2}}} = -1$$

Multiplying out and squaring the second and last terms give

$$\cos^2 r_2 \cos^2 i_1 = 1 - n^2 \sin^2 r_2 - \frac{\sin^2 i_1}{n^2} + \sin^2 r_2 \sin^2 i_1$$

By the use of $\cos^2 i_1 = 1 - \sin^2 i_1$ and a similar equation for $\cos^2 r_2$, one readily finds that this reduces to

$$(n^2 - 1) \sin^2 r_2 = (n^2 - 1) \sin^2 r_1$$

or

$$r_2 = r_1$$

Consequently, $i_2 = i_1$, showing that the light passage is symmetrical. In this case $r_2 = r_1 = A/2$ and $i_2 = i_1 = \frac{1}{2}(A + \delta)$, so that

$$n = \frac{\sin \frac{1}{2}(A + \delta)}{\sin \frac{1}{2}A}$$

giving a useful laboratory formula for the determination of the index of refraction of a prism from the angle of minimum deviation.

The variation of n with wave length is not represented exactly by any known formula. A useful approximation, containing only two constants, is the Cauchy formula

$$n = n_0 + \frac{B}{\lambda^2}$$

This can be made more exact by adding an inverse fourth power term with a constant multiplier. However, a better approximation, with three constants n_0 , λ_0 , and C , is the modified Hartmann formula

$$n = n_0 + \frac{C}{\lambda - \lambda_0}$$

Both of these are easily differentiated when it is desired to obtain a formula containing the rate of variation of n with wave length.

A prism spectrograph has a maximum aperture and minimum astigmatism when the prism is in the position of minimum deviation. Obviously, this condition cannot be simultaneously satisfied for all wave lengths. In instruments using a fixed prism, one can at best set the prism at the proper angle for the central portion of the spectrum.

In spectroscopes for visual observation it is common to use a combination of two refractions and one reflection to obtain a minimum deviation of usually 90° for all wave lengths in turn as the prism is rotated. The telescope may then be rigidly fixed at 90° with the collimator, and, as the prism is rotated by means of a

screw with a calibrated wave-length scale, each line in the center of the field has passed through the prism in a manner which satisfies the condition of minimum deviation. The **Pellin-Broca prism**, Fig. 7.18, is a type commonly employed. Although made in one piece, this prism is equivalent to two 30° prism components and one

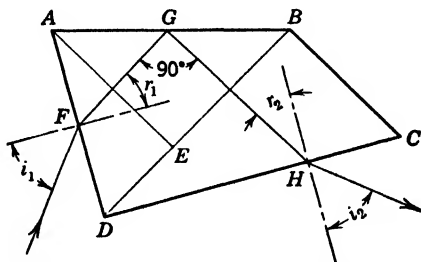


FIG. 7.18. Pellin-Broca constant deviation prism.

45° totally reflecting prism AEB . The deviation is a minimum when the rays FG and GH in the prism are respectively parallel to the bases of the two 30° prism components ADE and BDC . Since these rays are at right angles to each other under this condition, the corresponding external rays are also at right angles, the internal angles r_1 and r_2 being equal, making the angles i_1 and i_2 also equal.

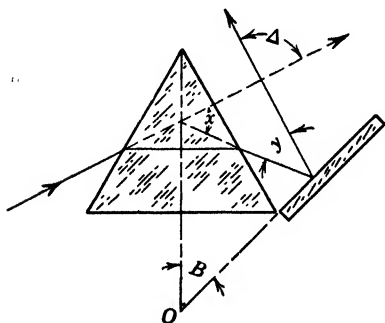


FIG. 7.19. Wadsworth mirror arrangement for constant deviation.

Another constant deviation device, diagrammed in Fig. 7.19, is a combination of a 60° prism and a mirror, known as a **Wadsworth mirror**. The angle of constant minimum deviation is equal to $180^\circ - 2B$, which is usually made equal to 90° . The proof that this arrangement gives a constant deviation of this amount is left

to the reader. The Wadsworth mirror is often employed with a rock-salt prism in one form of infrared spectrometer. The prism and mirror are rotated together by a calibrated screw, preferably about an axis passing through O .

Spectrum lines are naturally curved because of the greater deviation of skew rays than those travelling in a plane at right angles to the refracting edge. The curvature is approximately parabolic, the radius at the vertex being

$$R \doteq \frac{n^2 f}{2(n^2 - 1)} \tan i$$

where f is the focal length of the lens, and i is the angle of incidence. The curvature changes very slowly with wave length, so that it may be compensated by using a curved slit. This compensation, at either the exit slit or the entrance slit, is important in a **monochromator** or a spectroscope using a photocell or thermopile at the receiving end.

A spectroscopic instrument has four important characteristics, its **dispersion**, its **resolving power**, its **relative aperture**, and the **spectral range** which it transmits. The dispersion of a spectroscope is defined as the rate of variation of deviation with wave length

$$D \equiv \frac{d\delta}{d\lambda} = \frac{di_2}{d\lambda}$$

Sometimes **linear dispersion** is given in terms of centimeters per angstrom unit, and sometimes one states the **inverse dispersion** as so many angstrom units per centimeter or millimeter. The former is merely the angular dispersion D multiplied by the focal length of the objective lens and by the cosine of the angle of incidence on the receiving surface. The second is the reciprocal of this quantity. The magnitude of D is related to the angle of the prism and the variation of its index with wave length. The formula is readily derived by the following steps. Starting with

$$i_2 = \sin^{-1} (n \sin r_2)$$

one finds that

$$\frac{di_2}{d\lambda} = \frac{1}{\sqrt{1 - n^2 \sin^2 r_2}} \left(n \cos r_2 \frac{dr_2}{d\lambda} + \sin r_2 \frac{dn}{d\lambda} \right)$$

But

$$r_2 = A - r_1 = A - \sin^{-1} \frac{\sin i_1}{n}$$

therefore

$$\frac{di_2}{d\lambda} = \frac{1}{\sqrt{1 - n^2 \sin^2 r_2}} \left(\frac{n \cos r_2}{\sqrt{1 - \sin^2 r_1}} \frac{\sin i_1}{n^2} + \sin r_2 \right) \frac{dn}{d\lambda}$$

and

$$\frac{di_2}{d\lambda} = \frac{1}{\sqrt{1 - n^2 \sin^2 r_2}} \left(\frac{\cos r_2 \sin r_1}{\cos r_1} + \sin r_2 \right) \frac{dn}{d\lambda}$$

This reduces to

$$D = \frac{di_2}{d\lambda} = \frac{\sin A}{\cos i_2 \cos r_1} \frac{dn}{d\lambda}$$

since

$$\cos i_2 = \sqrt{1 - n^2 \sin^2 r_2}$$

At minimum deviation, $r_1 = r_2 = A/2$. Then

$$D = \frac{2 \sin A/2}{\sqrt{1 - n^2 \sin^2 A/2}} \frac{dn}{d\lambda}$$

If a 60° prism is used, $\sin A/2 = 1/2$, so that

$$D = \frac{1}{\sqrt{1 - n^2/4}} \frac{dn}{d\lambda}$$

Using Hartmann's dispersion formula for n , the expression for dispersion becomes

$$D = \frac{C}{(\lambda - \lambda_0)^2 \sqrt{1 - n^2/4}}$$

The **resolving power** R of a spectroscope is defined as the ratio of the wave length λ to the smallest change in wave length that can be resolved $\Delta\lambda_m$. Thus

$$R = \frac{\lambda}{\Delta\lambda_m}$$

If we let $\Delta\theta_m$ represent the smallest angle that can be resolved in the image space, one may write

$$\Delta\lambda_m = \frac{\Delta\theta_m}{D}$$

so that

$$R = D \frac{\lambda}{\Delta\theta_m}$$

In practice, the value of $\Delta\theta_m$ depends on slit width and residual aberrations as well as on diffraction, but the first two may be made so small that the resolving power of a properly designed spectroscope is finally limited by diffraction.

In Art. 11.2 it will be proved that, for a rectangular aperture of width a , the value of $\Delta\theta_m = \lambda/a$. It follows then that

$$R = Da$$

This general formula may be specialized for a prism spectroscope by noting, Fig. 7.20, that

$$a = L \cos i_2$$

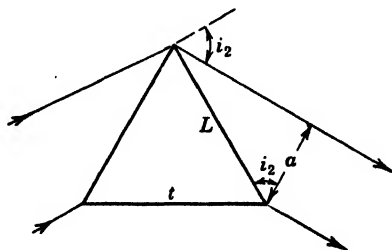


FIG. 7.20. The resolving power of a prism depends on the aperture width a .

and by using the previously derived expression for D . Thus it is found that

$$R = \frac{L \sin A}{\cos r_1} \frac{dn}{d\lambda}$$

At minimum deviation $r_1 = A/2$, so that this reduces to

$$R = 2L \sin \frac{A}{2} \frac{dn}{d\lambda} = t \frac{dn}{d\lambda}$$

Hence the resolving power of a prism depends on the rate of variation of n with wave length, i.e., the material used, and on the thickness t of the prism at its base. If several prisms are used in series without reduction in aperture, one may add their thicknesses. If the full aperture of a prism is not used, one should replace t by the difference, $t_2 - t_1$, between the lengths of the boundary rays in the prism.

In order to obtain the theoretical resolving power, the slit width must be negligible or infinitesimal. In this case the amount of light entering the spectroscope is also infinitesimal. Hence a compromise is obviously necessary. Sir Arthur Schuster was the first to discuss the resolving power of spectroscopes having a finite slit width. His results are expressed in terms of a "normal slit width" w_n , where

$$w_n = \frac{f_c \lambda}{8\rho_c}$$

in which f_c and ρ_c are respectively the focal length and exit-pupil radius of the collimator lens. In the case of the normal slit width, the optical paths from the various points on a line across the slit do not differ by more than $\lambda/4$. This tolerance agrees with interferometer tests of optical systems, which show that such changes in light path scarcely affect the quality of an image. The magnitude of the maximum intensity across a spectral line is given in terms of unity for the normal slit width in column 3, Table 7.1, and in terms of unity for an "infinitely" wide slit which gives a broad, flat-topped "line," in column 4. The **purity factor**, P , is

TABLE 7.1

Slit Width in Multiples of w_n	Purity Factor	Intensity	
		Unity for w_n	Unity for $w = \infty$
0.0	1.000	0.000	0.000
1.0	0.986	1.000	0.246
2.0	0.943	1.902	0.467
7.2	0.495	3.670	0.902
Infinity	0.000	4.069	1.000

the ratio of the practical resolving power with the finite slit width to the ultimate resolving power derived from diffraction theory for an infinitesimal slit.

It is commonly recommended that one select a spectroscope giving at least twice the resolving power actually needed, so that the slit may be opened to 7.2 normal. This reduces the resolving power to half its maximum and gives an intensity which is 90 per cent of its maximum. However, when the light source has a high enough intensity, it may be better to use the normal slit width, obtaining nearly (98.6 per cent) the theoretical resolving power with about one fourth of the maximum intensity.

Schuster's investigation has been extended by Van Cittert to include light sources which are small relative to their distance from the slit of the spectroscope. The results are summarized by Sawyer in *Experimental Spectroscopy*, pages 105-109.

PROBLEMS

1. The near and far points for an unaided eye are respectively at 15 cm and 60 cm. (a) What power spectacle lens should be prescribed? (b) What is the shortest distance of distinct vision with this lens?

2. If the near and far points of an unaided eye are respectively at 15 cm and 90 cm, what will be the range of distinct vision with a -1.0 -diopter spectacle lens?

3. A farsighted person cannot see objects clearly when they are nearer than 1 meter. What is the lowest power of a lens which will enable him to see clearly at the normal reading distance?

4. A set of trial lenses having a combined vertex power of $+10.0$ diopters is found to give satisfactory vision to an aphakic eye (no crystalline lens) when the posterior vertex of the set is 18 mm in front of the cornea. What vertex power will be required in the prescription lens if it is to be worn at a distance of 14 mm from the cornea?

5. A supplementary lens of $+2.0$ diopters is placed directly in front of a camera lens. If the object to be photographed is 40 cm away, for what distance should one set the focussing scale of the camera?

6. A portrait attachment consists of a $+2.0$ -diopter lens just in front of the camera objective. If the focussing scale is set for a distance of 3.5 ft, at what distance will objects be in sharp focus with the attachment in place?

7. Two lenses of 4-cm focal length are placed 3 cm apart on a common axis, forming a Ramsden eyepiece. From what point should incident rays diverge in order that the emerging rays be parallel?

8. Two lenses of a Huygens eyepiece have focal lengths of 3 cm and 1 cm respectively and are 2 cm apart on a common axis. If the final image formed by the second lens is at infinity, how far is the focus of the incoming rays from the first lens?

9. A telephoto camera lens has a converging lens of 10-cm focal length, and 8 cm behind it is a diverging lens of 3-cm focal length. (a) What is the distance behind the diverging lens of the image when the object is at infinity? (b) What is the equivalent focal length of the combination?

10. The focal length of the eyepiece of a compound microscope is 1.5 cm, and the tube length is 16 cm. What focal length objective should be used to obtain a magnifying power of 1000 diameters?

11. The focal lengths of the objective and eyepiece of a compound microscope are respectively 3 mm and 20 mm. The tube length is 160 mm. Compute the magnifying power of (a) the eyepiece, (b) the objective, and (c) the compound microscope.

12. A motion-picture projector using a film with a picture width of 24 mm throws an image 20 ft wide on a screen 100 ft away from the projection lens. What is the focal length of the lens?

13. An opera glass is to be constructed having a tube length of 8 cm and a magnifying power of four. What must be the focal lengths of the objective and eyepiece lenses?

14. A telescope has an objective of 10-in. focal length and an eyepiece of 1.5-in. focal length. (a) What is its magnifying power? (b) What is the actual size of the final virtual image of an object 10 ft high at a distance of 1000 ft if the final image is 20 in. from the eyepiece lens?

15. A telescope has an objective of 50-cm focal length. How much motion must be allowed the eyepiece if the telescope is to focus on objects as close as 5 meters as well as on objects very far away?

16. A microscope focussed on a mark must be raised 3.5 mm when a plate of glass 1 cm thick is placed over the mark and the microscope is refocussed on the image of the mark. What is the index of refraction of the glass?

17. Prove that the index of refraction of a prism of angle A is equal to $\sin(A + \delta)/\sin A$, where δ is the deviation obtained when the incident light is travelling normal to the first face or, conversely, emerging normal to the second face.

18. Find the largest angle of a prism made of material whose index of refraction is 1.65 which can transmit light at minimum deviation.

19. A light beam passes through a prism of crown glass, $n = 1.53$, having a 6° refracting angle. (a) What is the angular deviation? (b) Through what distance will the light be deflected transversely at a distance of 1 meter from the prism?

20. What is the angle of a flint-glass prism whose index of refraction is 1.695 which must be combined with a 5° crown-glass prism whose index is 1.523 in order that the angle of deviation may be zero?

21. It is desired to combine a flint-glass prism and a 10° crown-glass prism with bases in opposite directions so as to annul the dispersion between the C and F lines. If the following refractive indices are used, what angle should the flint-glass prism have? What will be the resulting deviation for the D lines?

INDEX OF REFRACTION	C LINE	D LINE	F LINE
Crown glass	1.5146	1.5171	1.5233
Flint glass	1.6224	1.6273	1.6394

22. If a ray of light is passing from glass into water, what is the greatest possible angle of incidence, the refractive indices of the glass and of water being 1.533 and 1.333 respectively?

23. The mean separation of a group of spectral lines at 2500 \AA is 0.25 \AA . Give the pertinent specifications that a prism spectrograph must meet in order to be suitable for photographing the resolved pattern.

24. A telescope is constructed of the following elements:

ELEMENT	FOCAL LENGTH	DIAMETER	SEPARATION
Objective	+300 mm	25.0 mm	
Air	302.4 mm
Field lens	+12 mm	12 mm	
Air	9 mm
Eye lens	+12 mm	8 mm	

The telescope is focussed for infinity. Compute the

- magnifying power of the telescope,
- size and location of the exit pupil,
- angular field of view in the object space,
- resolving power in seconds of arc. (Refer to Art. 11.4.)

Chapter 8. Stereoscopy

8.1 Binocular Vision

One may obtain a monocular sense of depth from perspective, shading, knowledge of absolute size, accommodation-convergence relation, and parallax. True **stereoscopy**, however, depends on a binocular sense arising from two slightly different images seen by the two eyes, Fig. 8.1. It arises from a mental interpretation of these two images which is based on experience. This true stereoscopic effect vanishes if the angle subtended by the eyes from the object point is less than about 30 seconds. Since the separation of the eyes is, on the average, 62 mm, the radius of stereoscopic vision is about 422 meters. With binoculars, this radius is increased because of two factors. In the first place, if the magnification is M , the radius is M times as large, since all parallax angles are increased in this ratio. For the same reason, if the entrance pupils of the binoculars have a greater separation d than the pupils of the eyes d_0 , the radius of stereoscopic vision R is d/d_0 times as great as with the unaided eyes. Thus we see that

$$R = R_0 M \frac{d}{d_0}$$

For example, if $M = 6$, and $d/d_0 = 2$, as with ordinary prism binoculars, then $R = 5000$ meters.

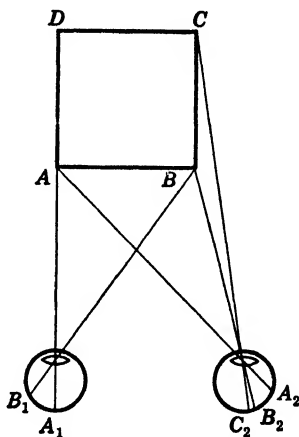


FIG. 8.1. Binocular vision.

8.2 Stereoscopic Camera and Viewer

Since stereoscopic vision is of great help in the examination of many objects, various visual instruments have been designed so

that each eye may see a suitably enlarged image from a different viewing point. Likewise, photographs taken by two lenses which are laterally displaced may be examined binocularly by means of a stereoscope to re-create the third dimension. In order to understand the principles on which these instruments are based, the stereoscopic camera and viewer may be taken as examples.

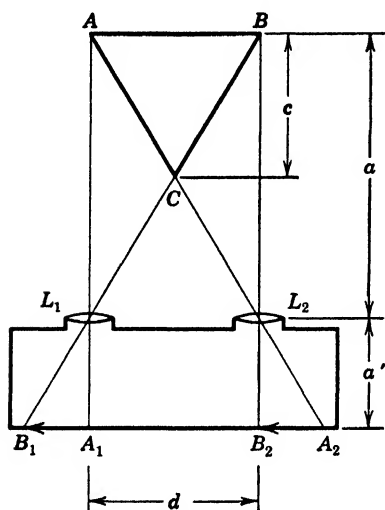


FIG. 8.2. Images formed in a stereoscopic camera.

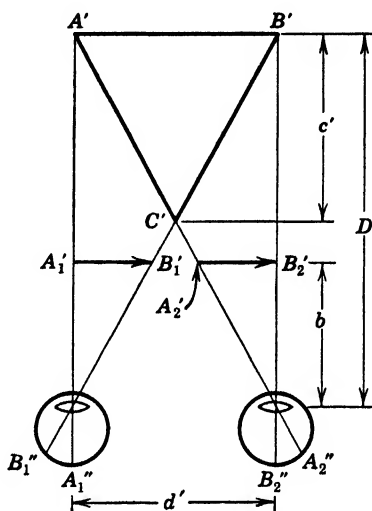


FIG. 8.3. Relation of space image $A'B'C'$ to the stereograms $A_1'B_1'$ and $A_2'B_2'$ which are viewed binocularly with a stereoscopic viewer.

Let L_1 and L_2 , Fig. 8.2, represent the two lenses of a stereoscopic camera which forms images of a pyramid ABC with A and B on the respective axes of L_1 and L_2 for simplicity. Then

$$\frac{A_2B_2}{AB} = \frac{A_1B_1}{AB} = \frac{a'}{a}$$

Consider the two enlargements $A_1'B_1'$ and $A_2'B_2'$ of A_1B_1 and A_2B_2 respectively, to be viewed by means of a stereoscopic viewer as shown in Fig. 8.3. The images on the retina will correspond to a pyramid $A'B'C'$ and will be interpreted as a three-dimensional object having the proportions indicated. Let the magnification or degree of enlargement in the viewer be

$$\frac{A_1'B_1'}{A_1B_1} = \frac{A_2'B_2'}{A_2B_2} = m$$

The distance D at which the base appears to be is

$$D = b \frac{A'B'}{A_2'B_2'} = b \frac{A_2B_2}{A_2'B_2'} \frac{A'B'}{A_2B_2} = \frac{b}{m} \frac{d'}{(a'/a)d} = a \left(\frac{1}{m} \right) \left(\frac{b}{a'} \right) \left(\frac{d'}{d} \right)$$

To obtain the same spatial relation of C' relative to $A'B'$ as for C relative to AB , the following relation must hold

$$\frac{c'}{A'B'} = \frac{c}{AB}$$

From the two figures one finds, however, that the two sides of the preceding equation are not equal in general, but have the ratio

$$\frac{c'/A'B'}{c/AB} = \frac{b/A_2'B_2'}{a'/A_2B_2} = \frac{b}{a'm} \equiv \frac{1}{M}$$

where M is called the apparent magnification since it refers to the relative sizes of the (tangents of) angles subtended by objects at the lenses of the eye and the lenses of the camera.

For orthostereoscopy the apparent depth and the apparent distance must respectively equal the actual depth and the actual distance, so that

$$\frac{c'}{A'B'} = \frac{c}{AB}$$

and

$$D = a$$

The first equation requires that

$$M = 1 \quad \text{or} \quad b = a'm$$

and the second requires that

$$\frac{1}{m} \frac{b}{a'} \frac{d'}{d} = 1$$

or, if $b = a'm$, that

$$d' = d$$

These conditions must be satisfied in the construction of stereoscopic viewers if correct perspective is desired.

In measuring stereoscopic photographs, a "floating spot" is often employed. This is created by the binocular fusion of the images of a spot in the two image planes $A_1'B_1'$ and $A_2'B_2'$. By providing a micrometer control for one or both spots, one may make the fused image move forward or backward with respect to $A'B'$. The micrometer screw may be calibrated in terms of the distance c .

8.3 Vectographs

The central layer of a sheet of polaroid is a plastic containing colloidal dichroic crystals which are uniformly aligned so that every element of area produces an equally high degree of polarization (Chapter 12). Dichroic crystals polarize light by their property of absorbing components of light vibrating in one direction to a much greater degree than components vibrating at right angles, the latter being transmitted. Crystals of iodosulphate of quinine are generally employed. If such a layer is coated over a gelatin film carrying a photographic image, it is possible, by chemical means, to destroy the dichroic crystals to a greater or lesser extent corresponding to the density of the photographic image, which is bleached in the process. This chemical action leaves a polarizing sheet having a varying degree of polarizability over its surface. A faint image remains because of the absorption in the remaining dichroic crystals. When this sheet is viewed through a completely polarizing sheet, called an analyzer, one can, by rotating it, vary the density of the image seen. The density of the image is practically zero when the analyzer is "parallel" with respect to the polarized vibrations from the image, and at a maximum when the analyzer is "crossed" with the sheet. In the latter orientation, the image can be seen with maximum clarity. Such a reproduction is called a **vectograph**.¹ In other forms of vectograph, the direction of polarization varies over the surface, and in still another form, the degree of polarization varies with wave length in the visible range. Every vectograph is characterized by differences in polarization over its area. For maximum contrast it must be viewed through a polarizing filter rotated to some optimum orientation.

It is of interest in stereoscopy that one may superpose two vectographs with the planes of polarization of their dichroic crys-

¹ LAND, *Journal of the Optical Society of America*, **30**, 230 (1940).

tals at right angles to each other. In this case, a rotation of a single analyzer causes one image to increase in contrast while the other disappears, and vice versa. If the images are stereoscopic views, one may construct a viewer by using two sheets of Polaroid with their polarizing directions at right angles and oriented so that one eye sees only one image, the other being faded out, while the other eye sees the other image. In this way one satisfies the conditions for a true binocular stereoscopic effect. Such stereovectographs have also been made in color. They may be projected on a screen, but the screen must be coated with an aluminum paint to avoid the depolarization which accompanies ordinary diffuse reflection. The two vectographs may also be supported on an aluminum-coated paper base and viewed directly through the analyzing viewer.

PROBLEMS

1. If the pupils of the eyes are 6.65 cm apart, how far away does a point appear whose stereoscopic images are 6.05 cm apart at a distance of 30 cm from the eyes?

2. The axes of the two lenses of a stereoscopic camera are 10 cm apart, and the focal lengths of the lenses are 5 cm. If the images of moderately distant objects are magnified 10 diameters and viewed from a distance of 30 cm, what should be the focal lengths of the viewing lenses, and what will be the ratio of the apparent depth to the actual depth? What will be the ratio of the apparent distance to the actual distance?

3. (a) What is the radius of stereoscopic vision of an observer whose interpupillary distance is 65 mm and whose binocular sense is 30 seconds of arc?

(b) What will be the radius when a 6-power binocular is used with its entrance pupils 110 mm apart?

4. Stereograms are made with a camera having a lens of 5-in. focal length. Specify a lens and prism combination for viewing such stereograms placed so that corresponding points on an infinitely distant object are 80 mm apart. Let $d' = 65$ mm.

II. PHYSICAL OPTICS

Chapter 9. Wave Propagation

9.1 Nature of Light

In the study of geometrical optics, it makes scarcely any difference what theory of light is assumed, since the theory is based on four non-committal, fundamental principles which are drawn from experience. It was pointed out that these principles are not accurately true, so that they lead to an incomplete, idealized situation which, under some conditions, is far from the facts. Thus the experimenter or careful observer often encounters the phenomena of interference, diffraction, and polarization, which cannot be explained by the principles of geometrical optics. In order to explain them, one is forced to assume at least that light has many of the characteristics of transverse waves. For many purposes, the kind of transverse wave is immaterial, but a study of electromagnetic waves shows that light is a member of this family of radiations. Thus optics may be considered a branch of electromagnetic theory.

Phenomena which are basically exchanges of energy between radiation and matter require the application of the quantum theory for their explanation. The union of the quantum theory with the wave theory is by no means obvious, but is being accomplished by means of quantum mechanics. The complete treatment of this unified theory is beyond the scope of this book. All that we need to assume is the existence of electromagnetic waves, coupled with the idea of their quantized interaction with matter. Thus exchanges in energy take place by increments of amount $\pm h\nu$, while exchanges in momentum take place by increments of $\pm h/\lambda$, λ being the wave length, ν the frequency, and h Planck's constant of action.

9.2 Wave Theory

A wave is characterized by its amplitude s_0 , its period T or frequency ν , and its state of polarization. The equation for a simple harmonic plane wave may be written in the form:

$$s = s_0 \sin \left[\frac{2\pi}{T} \left(t - \frac{lx + my + nz}{V} \right) + \epsilon \right]$$

In practice one can never obtain one simple harmonic plane wave, but it can be closely approximated, and actual waves can always be represented as a sum of such waves. The symbols in the above equation are defined in the following paragraphs in connection with an analytical discussion of the equation. First of all, one notes that, if all quantities are kept constant except the time t , the equation evidently represents a simple harmonic vibration. The quantity s_0 is thus recognized as the maximum value of s or the **amplitude** of the vibration. Physically, s and s_0 may be identified as either the electric field strength, the dielectric displacement, or the magnetic field strength. The **dielectric displacement** is defined as the vector sum of the electric field strength and 4π times the **dielectric polarization**, the latter quantity being the electric dipole moment per cubic centimeter. Usually the dielectric displacement is chosen as the light vector s , for then the equations derived for light propagation inside a crystal have the same form as those first developed by Fresnel. In free space and in isotropic media, the dielectric displacement not only is proportional to the electric field strength but also is always in the same direction as the field. The constant of proportionality is called the **dielectric constant**. In these cases the equations for light propagation have the same form regardless of the physical interpretation of the light vector s . It is in such media that the waves are finally observed. Light propagation in crystalline media will not be considered until Chapter 12.

A plane wave of ordinary light may be resolved into a superposition of numerous elementary wave trains having the above form, with s_0 assuming all possible directions normal to the direction of propagation during small, successive time intervals. If, however, the vector s_0 remains constantly in one plane, the light is said to be *plane polarized*. The plane of polarization is defined as the plane perpendicular to the dielectric displacement. This

plane is being gradually replaced in optical literature by the vibration plane which contains the light vector. It should be noted that vibratory displacement, like any vector, may be resolved into two orthogonal components having an orientation most convenient for the problem being discussed. If the light is unpolarized, the net components will always be equal, whereas if it is polarized, they will be the sine and cosine components of the light vector.

The quantity T in the above equation is the **period**, or time for one complete cycle of change at any point. Its reciprocal is the **frequency**, or number of vibrations per second, and it is related to the wave length λ by the equation $\lambda = VT$, where V is the velocity of propagation of the wave. Since V depends on the transmitting medium whereas T does not, the wave length corresponding to light of any given color will also depend on the medium. Hence, whenever colors are designated by their wave lengths, the transmitting medium should also be specified. If this is not given, it may be assumed that the wave length is measured in free space. Since the velocity of light in air is only 0.03 per cent smaller than in free space, the wave length in air will likewise be that much smaller, and the difference, being small, may frequently be neglected. The visible spectrum extends from about 3800 Å to about 7600 Å in wave length, where $1 \text{ Å} = 10^{-8} \text{ cm}$. The human eye is most sensitive to a wave length of 5550 Å, which may be taken as the effective wave length of white light. For purposes of exact optical measurements, light of as nearly a single wave length as possible must be used. Such light may be obtained from sodium or mercury-vapor lamps with the interposition of filters to remove the undesired spectrum lines. The wave lengths used most frequently are mercury 5461 Å and sodium 5893 Å, a green and a yellow line, respectively.

The quantity enclosed in brackets in the expression on page 107 for a plane wave is called the **phase** of the wave. The locus of points in xyz space at which the phase is constant is called a **wave front**. In this case it is a plane

$$lx + my + nz = \text{constant}$$

the values of l , m , and n being the cosines of the angles made by the wave normal with the x , y , and z axes, respectively. The quantity V is the velocity of advance of the wave in the direction of the

wave normal, as can be readily seen by considering the ratio of the distance of propagation to the corresponding change in time for any given value of the phase. The phase constant ϵ is needed only when one must compare the phases of two or more waves, for otherwise one may merely change the time origin and incorporate ϵ within t .

When V is independent of the direction of the light vector s_0 , the medium is said to be **isotropic**. The wave fronts are then at right angles to the direction of energy flow which defines the light ray. The ordinary laws of refraction hold when light passes from one such medium into another. Moreover, the rays and wave normals always lie in the plane of incidence defined by the normal to the surface and the normal to the incident wave. The angles ϕ_1 and ϕ_2 , which the wave normals in the two media make with the normal to the surface, are related by the equation, Snell's law,

$$\frac{\sin \phi_1}{\sin \phi_2} = \frac{V_1}{V_2} = \text{constant}$$

The constant is called the **index of refraction** of the second medium with respect to the first and is equal to the ratio n_2/n_1 , where n_2 and n_1 are the indices of the two media with respect to a vacuum. Thus $n_2 = c/V_2$ and $n_1 = c/V_1$, where c is the speed of light in free space, and n_2 and n_1 are called the **absolute indices of refraction** of the respective media.

When V depends on the direction of the light vector, the medium is said to be **anisotropic**. Such media are crystals (with the exception of those belonging to the cubic system), all transparent media subjected to stresses, and colloidal materials with oriented fibrous or lamellar micelles. Their optical properties are discussed in Chapter 12.

9.3 Wave and Group Velocity

The velocity V of a wave is not measured directly by the usual methods for measuring the speed of light. In free space or air, one usually measures, instead of V , the signal or **group velocity** U of some impressed change in amplitude. To discover how U and V are related, assume a mathematically simple sort of modulation which may be represented by the sum of two waves having the same amplitude but slightly different frequency. These waves will

have slightly different phase velocities in a dispersing medium. The superposition of two such waves gives a beat wave, as shown in Fig. 9.1. Analytically,

$$s = A \left[\sin \frac{2\pi}{T_1} \left(t - \frac{x}{V_1} \right) + \sin \frac{2\pi}{T_2} \left(t - \frac{x}{V_2} \right) \right]$$

$$\doteq 2A \sin \frac{2\pi}{T} \left(t - \frac{x}{V} \right) \cos \left[t \left(\frac{1}{T_1} - \frac{1}{T_2} \right) - x \left(\frac{1}{V_1 T_1} - \frac{1}{V_2 T_2} \right) \right]$$

It is assumed that T_1 and T_2 , as well as $V_1 T_1$ and $V_2 T_2$, are nearly equal, so that one may write their average values as T and VT re-

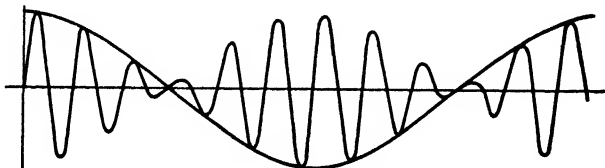


FIG. 9.1. A wave group.

spectively. The cosine factor represents the modulation of the sine wave. In Fig. 9.1, this is shown as the *envelope* of the higher-frequency "carrier wave." The velocity of the sine or the cosine wave is found by dividing the coefficient of the t term by the coefficient of the x term.¹ Thus for the carrier wave

$$v = \frac{1/T}{1/VT} = V$$

and for the group

$$U = \frac{\Delta(1/T)}{\Delta(1/VT)} \doteq \frac{d\nu}{d\bar{\nu}}$$

¹ Let

$$y = A \sin (at - bx)$$

Then

$$y = A \sin [a(t + \Delta t) - b(x + \Delta x)]$$

at a later time and farther along. The same y is found everywhere if

$$a \Delta t = b \Delta x$$

Hence

$$v = \frac{\Delta x}{\Delta t} = \frac{a}{b}$$

where ν is the frequency, and $\bar{\nu}$ is the wave number, $\bar{\nu} = 1/\lambda = 1/VT$. Since $\nu = V\bar{\nu}$, one readily obtains

$$U = V + \bar{\nu} \frac{dV}{d\bar{\nu}}$$

or, in terms of λ ,

$$U = V - \lambda \frac{dV}{d\lambda}$$

In the optical region, the velocity of light in transparent media increases with wave length, since for these media the index of refraction, $n = c/V$, diminishes. Hence the group velocity is less than the wave velocity by an amount depending on the rate of change of V with λ . If there is no dispersion, as in free space or (nearly) in air, then $U = V$, and the distinction between the two velocities disappears or becomes imperceptible. On the other hand, measurements which have been made of the speed of light U in water and particularly in carbon disulphide require an appreciable correction to obtain V , from which one may find the index of refraction c/V .

9.4 Measurement of the Speed of Light

In all experimental methods for measuring the speed of light except Bradley's method, which employs the aberration of light, one is forced to modulate the wave by changing its intensity and then measure the signal or group velocity. One excellent method is that of Foucault, as perfected by Michelson, which employs a rotating mirror as a timing device. Essentially, one measures the short time interval Δt taken by light to travel over a distance L and back again, by observing the angle through which a uniformly rotating mirror turns in the same time. If the mirror makes n revolutions per second and rotates through α radians in this time interval, then $\Delta t = \alpha/2\pi n$. Michelson used many-sided mirrors in his last experiments. For example, with an octagonal mirror, the position of the mirror when a wave train is leaving a face A toward the distant reflector is shown in Fig. 9.2. If the rate of rotation is adjusted properly, the returning wave train is reflected by face B after it has turned through exactly $\alpha = \pi/4$ to B' . The returning light would then be undeflected, as if the mirror were placed in the position shown and not rotated at all. In this de-

scription, the method and apparatus are greatly simplified. For details, one should refer to Michelson's papers or his book *Studies in Optics*. In his 1924 experiments the distance L was about 22 miles from Mount Wilson Observatory to a station on Mount San Antonio. The rate of rotation n was about 529 sec^{-1} , this being measured by comparison with a standard tuning fork. The weighted mean of 1763 observations was $299,796 \pm 4 \text{ km/sec}$ for the speed of light corrected from air to vacuum. Later, Michelson

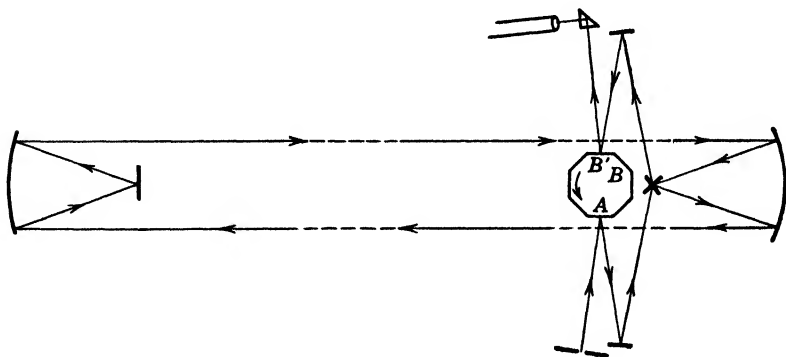


FIG. 9.2. Diagram of Michelson's rotating-mirror method for measuring the speed of light.

measured the speed directly in a mile-long vacuum tube in order to eliminate the uncertainties arising from the reduction of the results from air to vacuum. These measurements were completed after Michelson's death in 1931 by his associates, Pease and Pearson, who obtained the value

$$V = 299,774 \pm 1 \text{ km/sec}$$

This is so close to the value of the speed of electromagnetic waves according to Maxwell's theory, using the results of purely electrical experiments performed by Rosa and Dorsey,

$$c = 2.99791 \times 10^{10} \text{ cm/sec}$$

that light is now universally considered to be an electromagnetic radiation.

The speed of light was measured more recently by W. C. Anderson with considerable precision.² The method used is essentially

² W. C. ANDERSON, *Journal of the Optical Society of America*, **31**, 187 (1941)

a modernized version of Fizeau's method, which employed a toothed wheel to interrupt the wave train. A beam of light from the source Q , Fig. 9.3, is alternately transmitted and intercepted by a Kerr cell arrangement at K (see Art. 19.6) which is connected to an oscillating circuit. A sequence of modulated pulses of length L and spacing L are transmitted by the cell. These pulses are divided at the half-silvered mirror at H . One short path leads to

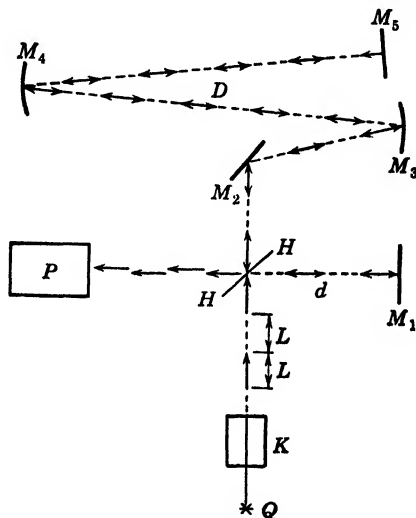


FIG. 9.3. Diagram of Anderson's Kerr-cell method for measuring the speed of light.

the mirror M_1 and back through the plate H to a photocell receiver and amplifier at P . The longer path leads to mirrors M_2 , M_3 , M_4 , M_5 and back to H , where part of this light is also reflected into the photocell. If the greater distance from H to M_5 is D and the shorter distance from H to M_1 is d , the pulses will arrive in opposition, giving a minimum response in the receiver, whenever $2(D - d)$ is some odd multiple of the pulse length L . If f is the frequency of the oscillator and U the group velocity of light, then

$$L = \frac{U}{2f}$$

and the minima occur when

$$2(D - d) = (2n + 1) \frac{U}{2f}$$

n being some integer. Since the speed of light is known quite accurately, the number n may be computed and rounded off to the nearest integer. Then a more accurate value of U may be found with the aid of the above equation. For a more complete discussion, as well as an improved variation of the procedure given above, the reader is referred to the original article. The final result given by Anderson for the speed of light in free space is

$$c = 2.99776 \pm 14 \times 10^{10} \text{ cm/sec}$$

PROBLEMS

1. One of Jupiter's satellites has a period of 42 hours, 28 minutes, 42 seconds. What will be the maximum period observed from the earth, which has a mean orbital speed of 18.5 miles per second? (The first measurement of the speed of light by Roemer was based on the difference in time between the observed and predicted eclipses of one of Jupiter's satellites.)

2. If light reflected from one face of a sixteen-sided mirror rotating at 400 rps returns from a distant mirror after the rotating mirror has turned through just one sixteenth of a revolution, how far is the distant mirror?

3. Show that one can eliminate the necessity of measuring the entire distance d in Anderson's method by replacing the mirror M_3 by a returning mirror M_5 , thus greatly reducing the distance D , and moving the mirror M_1 by a measured small distance sufficient to obtain a minimum. Assume that the two paths are made nearly equal by this reduction in D .

4. Compute the phase velocity and the group velocity for light of 5893 \AA wave length travelling in crown glass and in flint glass, using the indices given in problem 21, Chapter 7.

Chapter 10. Interference of Light

10.1 Young's Experiment

Thomas Young first showed, in 1800, that one could cause two beams of light to overlap in such a way that alternating regions of light and darkness are observed. He called this phenomenon **interference** between the two light beams. The apparatus needed is of the simplest sort. Light from a source passes through a narrow slit at S , Fig. 10.1, and then through two parallel slits S_1 and

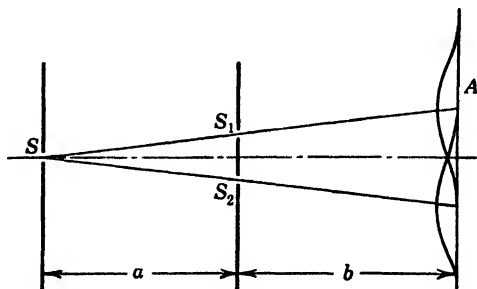


FIG. 10.1. Young's experiment, showing the overlapping of beams due to diffraction of light passing through the two narrow slits.

S_2 at a distance a from S . The illumination on a surface A at a distance b from the double slit shows the interference bands. The separation of S_1 and S_2 should be quite small compared with the distance b , and the width of S should be very small compared with a . Unless the source is very bright, it is difficult to see the bands, but they may be readily recorded on a photographic plate at A (see Fig. 10.2), or one may remove A and observe the bands with the aid of an eyepiece or magnifier.

As a wave front from S reaches the double slit, two nearby portions pass through the respective slits. A considerable amount of diffraction, Art. 11.2, due to the narrowness of the slits, causes the light to be distributed over an appreciable area of A . Each slit by itself produces a characteristic diffraction pattern of the type

to be discussed in the next chapter. However, the illumination due to both beams is not the sum of the two, but is found to be very low at some locations and to exceed the sum at others. These phenomena are called respectively **destructive interference** and **constructive interference**. Fundamentally, there is a re-distribution of the light, the intensity at each point being propor-

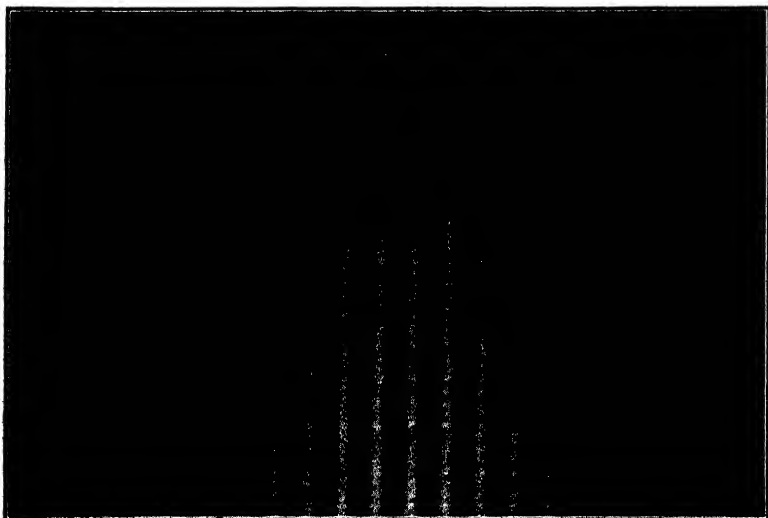


FIG. 10.2. Young's experiment. Interference bands in mercury green light, using a double slit made by ruling two parallel scratches through the backing of a mirror with a knife blade.

tional to the square of the amplitude of the wave resulting from the addition of the two waves from S_1 and S_2 respectively.

10.2 Addition of Waves

The two waves superpose, each one travelling as if the other one were not there, so that

$$s = s_1 + s_2 = A_1 \sin \left[\frac{2\pi}{T} \left(t - \frac{d_1}{V} \right) + \epsilon_1 \right] \\ + A_2 \sin \left[\frac{2\pi}{T} \left(t - \frac{d_2}{V} \right) + \epsilon_2 \right]$$

d_1 and d_2 being the respective distances from S_1 and S_2 to some point P , the other symbols being as previously defined, Art. 9.2.

The phase constants ϵ_1 and ϵ_2 in Young's experiment will be $2\pi/\lambda$ times the distances SS_1 and SS_2 . By expanding the sine functions and collecting terms, one obtains a result which may be written in the form

$$s = A \sin \left(\frac{2\pi}{T} t - \delta \right)$$

where

$$A^2 = A_1^2 + A_2^2 + 2A_1A_2 \cos \Delta$$

in which Δ is the phase difference

$$\Delta = \frac{2\pi}{TV} (d_1 - d_2) + (\epsilon_2 - \epsilon_1) = \frac{2\pi}{\lambda} (d_1 - d_2) + (\epsilon_2 - \epsilon_1)$$

When the two waves are from different sources, or even from different points in the same source, the phase constants ϵ_1 and ϵ_2 change very frequently by arbitrary amounts and the waves are said to be **non-coherent**. In this case, one can only observe the time average of A^2 which is equal to $A_1^2 + A_2^2$. Since the intensities of all kinds of waves are proportional to the squares of their amplitudes, this result shows that the resultant intensity is just the sum of the two individual intensities. It will be convenient to consider the constant of proportionality between the square of the amplitude and the intensity to be unity. It will have this value in some appropriate units.

If the cross-product in the expression for A^2 is not to average out, the two waves must preserve a constant phase difference at all times, that is to say, they must be **coherent**. In the case of light, such coherent waves are obtained only by the division of a single wave into two parts, which are reunited after travelling along paths which differ by an amount such as $n_1d_1 - n_2d_2$. The difference between the phase constants ϵ_1 and ϵ_2 will be either 0 or π in practically all cases in which we will be interested. The union of coherent waves will produce amplitudes, and therefore intensities, which depend greatly, and in a characteristic manner, on the value of Δ . In particular, the intensity will be greatest when Δ is some integral multiple of 2π , and it will in this case be even greater than the sum of the two separate intensities, since

$$A^2 = A_1^2 + A_2^2 + 2A_1A_2$$

In such a case of constructive interference the resulting amplitude

is simply the sum of A_1 and A_2 . On the other hand, when $\Delta = (m + \frac{1}{2})2\pi$, with $m = 0, 1, 2, 3 \dots$, one finds that

$$A^2 = A_1^2 + A_2^2 - 2A_1A_2$$

the amplitude being equal to the difference between A_1 and A_2 . Thus the resultant intensity is less than the sum of the separate intensities. The average intensity in the two cases is, however, still equal to the sum of the two intensities. Interference phenomena always contain regions in which there is an excess of energy and others where there is a deficiency, but of such an amount that

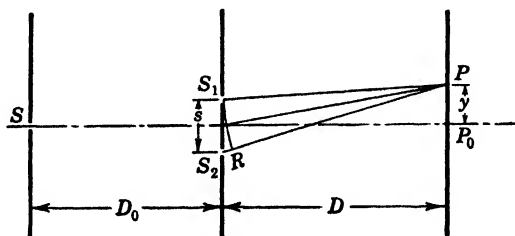


FIG. 10.3. Double slit, showing the path difference from the two slits S_1 and S_2 to a point P .

the total energy is conserved. No absorption of light is necessarily associated with interference.

The specific cases given above are only those for which the intensity has a maximum or a minimum value. Intermediate results are included in the formula for A^2 , the value of which may be found graphically by noting its similarity to the cosine law of trigonometry. One may accordingly construct an **amplitude triangle** with a side of length A_2 making an exterior angle Δ with a side of length A_1 . The closing side of length A is the resultant amplitude.

In Young's experiment, Fig. 10.3, the difference in path d from S_2 and S_1 to P is given by

$$d = S_2P - S_1P$$

The arc S_1R is always small and may be replaced by its chord. Moreover, the angle S_2S_1R is so small that one may take the angle as equal to its sine or tangent, so that

$$\frac{d}{s} = \frac{y}{D}$$

The difference in phase between the waves arriving at P from S_1 and S_2 is accordingly

$$\Delta = \frac{2\pi d}{\lambda} = \frac{2\pi sy}{\lambda D}$$

assuming that the slits S_1 and S_2 are equally distant from S so that $\epsilon_1 - \epsilon_2 = 0$. The locations of maxima of intensity will correspond to values of $\Delta = m 2\pi$, or $d = m\lambda$. Thus for the maxima

$$y = m \left(\frac{D\lambda}{s} \right)$$

and for the minima

$$y = \left(m + \frac{1}{2} \right) \left(\frac{D\lambda}{s} \right)$$

where m is an integer called the **order of interference**.

Since λ is very small for visible light, it is necessary to make s much smaller than D in order to resolve the bands. For distinct interference bands, S must be narrow enough so that the difference in phase of waves arriving at S_1 and S_2 from its various parts will not exceed about half a wave length. The slits S_1 and S_2 must be narrow enough so that the two central diffraction maxima overlap each other by at least two interference bands. Since the central diffraction maximum for a single slit has a total width of $2D\lambda/e$, where e is the slit width of S_1 and S_2 , this means that s must be more than three times as large as e .

10.3 Fresnel's Mirrors and Biprism

The necessarily narrow slits in Young's experiment make the intensity of the pattern quite low. Moreover, the interference phenomenon is dependent on the overlapping of waves due to diffraction at the double slit. In order to overcome these limitations, Fresnel devised his double mirror and biprism. The underlying theory is quite similar to that of Young's experiment: one must only derive an expression for s , the separation of the double images of S , which are now obtained by reflection or refraction, as shown in Figs. 10.4 and 10.5. If α represents the angle between the two mirrors, Fig. 10.4, the angle subtended at the intersection of the two mirrors by the two images S_1 and S_2 of S will be 2α , as can be readily proved by applying the law of reflection. The images and

source lie on a common circle of radius D_1 . The resulting optical arrangement is that of two coherent sources which are separated by a small distance $s = 2aD_1$, from which light is sent to points such as P at a distance $D_1 + D_2$. In essential respects this is

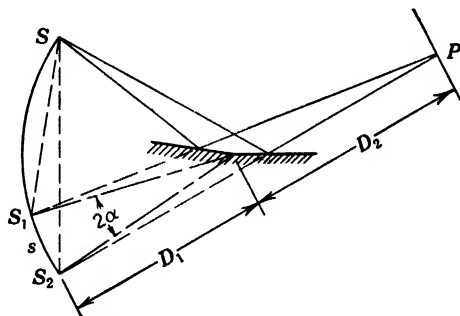


FIG. 10.4. Fresnel mirrors, showing formation of coherent images of the slit S . similar to Young's experiment. Hence a suitable alteration of the formulas to fit the present case gives for the location of the interference maxima

$$y = m \frac{(D_1 + D_2)\lambda}{2D_1\alpha}$$

and for the minima

$$y = \left(m + \frac{1}{2}\right) \frac{(D_1 + D_2)\lambda}{2D_1\alpha}$$

If the slit S is moved close to the plane of a *single* mirror, one can observe interference between the waves from S and from its

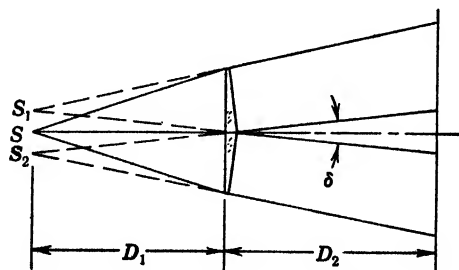


FIG. 10.5. Fresnel biprism, showing formation of coherent images of the slit S . reflected image. Such an arrangement is called a **Lloyd's mirror**.

In the Fresnel biprism, Fig. 10.5, the lower half of each wave from S which passes through the prism is deviated through a small

angle toward the upper half. It has been shown that, for prisms of small angle α and index n , the deviation equals $(n - 1)\alpha$. The distance between the coherent images S_1 and S_2 will therefore be

$$s = \delta D_1 = 2D_1(n - 1)\alpha$$

Thus one finds that the intensity maxima are located at values of y given by

$$y = \frac{m(D_1 + D_2)\lambda}{2D_1(n - 1)\alpha}$$

A photograph of a biprism interference pattern is shown in Fig. 10.6.

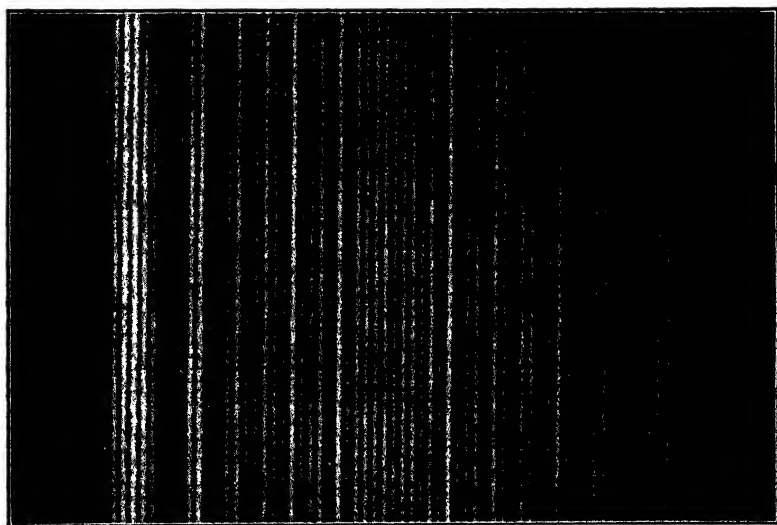


FIG. 10.6. Interference bands obtained with a Fresnel biprism. The superposed intensity fluctuations are due to diffraction arising from the limitation of the wave front at the apex of the biprism. Compare with Fig. 11.16.

10.4 Michelson Stellar Interferometer

This interferometer has the mirrors M_1 , M_2 , M_3 , and M_4 , Fig. 10.7, supported on a rigid beam. The separation s of mirrors M_1 and M_4 may be increased up to about 20 ft. The mirrors M_2 and M_3 reflect the light into the telescope. If the mirrors M_1 and M_4 are relatively close together, interference bands are observed crossing the diffraction disk of the star in the focal plane of the

telescope at E . One has essentially a double aperture M_1M_4 as in Young's experiment, the star serving as the source S . The paths from the different elements of area of the star to the two mirrors of the double aperture vary by less than $\lambda/2$ when the separation

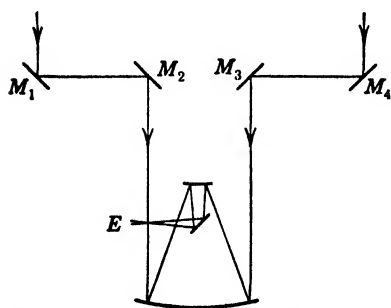


FIG. 10.7. Light path in Michelson's stellar interferometer.

s is small, so that the fringes are quite distinct. However, as one increases s , the path difference from the elements of area of the star to M_1 and M_4 finally becomes great enough so that the bands are obliterated. To understand better the conditions for this disappearance, consider, for example, a rectangular source of width w at a distance D , Fig. 10.8. Let the optical

paths from the edge A to M_1 and M_4 be equal. Then the optical paths from B to M_1 and M_4 will differ, see Fig. 10.8(a), by the amount

$$d = s \frac{w}{D}$$

If $d = \lambda$, the source may be divided into halves, as in Fig. 10.8(b), and in each half there will be corresponding strips for which the

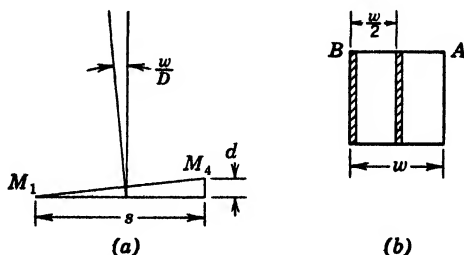


FIG. 10.8. (a) Path difference of light from off-axis portion of a distant object. (b) Elements of area of distant object whose width is w .

difference in path is $\lambda/2$, so that the fringes at E produced by light from these strips will be complementary and hence will obliterate each other. A second disappearance will occur when $d = 2\lambda$, but

usually the star is small enough so that only the first disappearance is observed. In this case

$$\lambda = s \frac{w}{D}$$

and the angle α subtended by the "star" will be

$$\alpha = \frac{w}{D} = \frac{\lambda}{s}$$

For a circular disk of uniform intensity, it can be shown that a factor of 1.22 should be included, giving

$$\alpha = 1.22 \frac{\lambda}{s}$$

In the case of the star **Betelgeuse**, the bands were found to disappear at $s = 306.5$ cm, although fringes could still be observed in the images of neighboring stars. Assuming an effective wave length of 5.75×10^{-5} cm, this gives for α a value of 0.047 second of arc. Since the parallax of Betelgeuse is 0.018 second, its linear diameter w is 240×10^6 miles, or nearly as large as the orbit of Mars. Other stars have since been measured by the same method.

10.5 Rayleigh Gas Interferometer

This is a modification of Young's experiment, applied to the measurement of the indices of refraction of gases or of very dilute solutions. Light from the source S , Fig. 10.9, is collimated by the

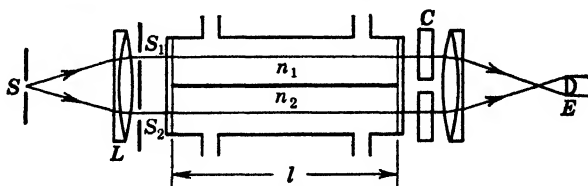


FIG. 10.9. Light paths in a Rayleigh gas interferometer.

lens L , passing through the double slit S_1S_2 and then through two adjacent tubes containing the two gases to be compared, the indices of which are represented by n_1 and n_2 , the tube length being l . The interference bands are observed with the aid of a cylindrical

lens at E . A cylindrical lens is preferred, since it magnifies the interval between the bands without diminishing the intensity as much as a spherical lens, which would also increase the spread of the light along the length of the bands. The difference in optical path through the two tubes is $(n_2 - n_1)l$, and this causes a fringe shift of m spaces, where

$$m = (n_2 - n_1) \frac{l}{\lambda}$$

The fringe shift may be counted as one gradually introduces the gas whose index is to be measured, but usually one makes measurements by compensating for the difference in path by rotating the compensating plate C . It can be proved that the increase in path d produced by rotating a compensator, whose index is n and thickness is t , through an angle i from the perpendicular is given by

$$d = t(\sqrt{n^2 - \sin^2 i} - \cos i - n + 1)$$

The fringe shift due to the difference between n_2 and n_1 may thus be compensated by rotating the plate C through some angle by means of a graduated micrometer screw. White light fringes are used which have one unique achromatic band at the center of the pattern. This achromatic band is brought back to its "zero" position by the calibrated compensator. One may use this interferometer for many kinds of analyses of gases, a commercial form by Zeiss being available for this purpose. In this form, the light is doubled back on itself, the entrance slit being just to one side of the eyepiece. The light is reflected into S from the side by means of a small totally reflecting prism. The instrument is built rigid enough so that it is portable.

10.6 Interference in Films and Plates

At any interface between two media, light is partially reflected and partially transmitted. When such waves are suitably reunited, we have a type of interference which is classified as due to **division of amplitude**. Newton's rings and similar phenomena due to partial reflections at two successive interfaces are familiar examples of this kind of interference.

The ratio of the amplitudes R and E of the reflected and incident waves at perpendicular incidence is given by Fresnel's reflection formula, Art. 12.3,

$$\frac{R}{E} = - \frac{n_2 - n_1}{n_2 + n_1}$$

where n_1 is the index of the first medium, and n_2 the index of the second. The Fresnel formulas for the general case of oblique reflection are derived and discussed in Art. 12.3. The negative sign, which applies even to oblique reflection up to the polarizing angle, indicates a 180° change in phase at the interface when n_2 is greater than n_1 . This must be added to the phase change due to the difference in optical path. The formula also shows that the magnitude of the amplitude of a normally reflected wave is one fifth that of the incident wave when the light is reflected either externally or internally at an interface between air and a material having an index of refraction of 1.5, such as crown glass. At external reflection there is a phase reversal, while at internal reflection there is none.

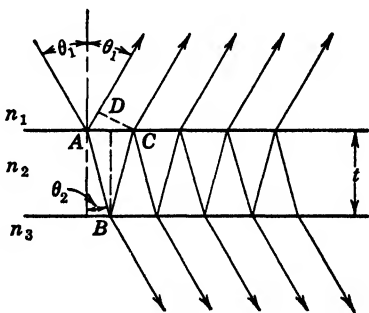


FIG. 10.10. Multiple reflections in a plane-parallel plate.

Let t be the thickness of a film of material whose index of refraction is n_2 , and let n_1 and n_3 be the indices of the media on its respective sides, Fig. 10.10. The optical path difference between any two of the successive reflections shown in the figure will be given by

$$\begin{aligned} d &= n_2(\overline{AB} + \overline{BC}) - n_1\overline{AD} \\ &= \frac{2n_2t}{\cos \theta_2} - n_1 2t \tan \theta_2 \sin \theta_1 \end{aligned}$$

Using Snell's law, $n_1 \sin \theta_1 = n_2 \sin \theta_2$, it is found that

$$d = \frac{2n_2t}{\cos \theta_2} (1 - \sin^2 \theta_2) = 2n_2t \cos \theta_2$$

Note that the difference in path is greatest for perpendicular incidence, so that it decreases as the angle of incidence is increased. If the index n_2 is either less than or greater than both of the indices

n_1 and n_3 , the phase difference between the first two of the successive reflections is

$$\Delta = \frac{2\pi}{\lambda} 2n_2t \cos \theta_2 \pm \pi$$

The other successive multiple reflections occur without the phase reversal π . To find the actual magnitudes of the intensity values one must combine all the amplitudes. However, since the amplitudes of the multiple waves beyond the second usually diminish very rapidly, one may find the conditions for maxima and minima without considering them. They effectively alter the amplitude of the second wave by a small amount which does not alter the conclusions given here.

Since there is destructive interference whenever Δ is any odd multiple of π , the first term on the right-hand side of the above equation, i.e., the retardation due to path difference, must in this case be an even multiple of π . Hence we find a minimum intensity when

$$2n_2t \cos \theta_2 = m\lambda, \quad m = 0, 1, 2, 3, \dots$$

and a maximum intensity when

$$\Delta = m2\pi, \quad m = 0, 1, 2, 3, \dots$$

or

$$2n_2t \cos \theta_2 = (m + \frac{1}{2})\lambda$$

On the other hand, if $n_1 > n_2 > n_3$, or if $n_1 < n_2 < n_3$, there is no extra π to be included in the expression for Δ , so that a *maximum* intensity occurs when

$$2n_2t \cos \theta_2 = m\lambda$$

and a *minimum* intensity when

$$2n_2t \cos \theta_2 = (m + \frac{1}{2})\lambda$$

The first set of conditions, for example, applies to a soap film in air or an air film between glass plates (Newton's rings). The second set applies to a fluoride film on glass, such as is often used to reduce losses due to reflection, the best thickness being $t = \lambda/4n_2$. For most perfect interference the reflectances at the two surfaces of the film must be equal. Fresnel's formula shows that this is the case if $n_2 = \sqrt{n_3n_1}$, which becomes equal to $\sqrt{n_3}$ if $n_1 = 1$.

The conditions for maximum and minimum intensity in transmitted light are readily seen to be the reverse of those applying to reflected light in all cases. Thus a maximum transmittance corresponds to a minimum reflectance, and vice versa. This was first observed by Newton and is now applied to coatings on lenses.

The colors of thin films are due to selective interference arising from a variation in phase difference with wave length. The wave lengths at which maxima or minima occur, depending on the relative indices, are determined by the formula

$$\lambda = \frac{2n_2t \cos \theta_2}{m}; \quad m = 0, 1, 2, 3, \dots$$

which indicates a sequence of bands in the spectrum. When the numerator is large compared with the wave length of visible light, the values of m needed to observe visible effects must also be large. In this case there will be a large number of fluctuations in the visible spectrum (channelled spectrum). With a sufficiently great number of maxima, no interference colors are seen, since the relative intensities of the primary red, green, and blue will not be altered. The most brilliant interference colors are observed when the spectrum contains only one or two interference maxima. The sequence of colors observed with increasing path difference is known as **Newton's color scale**. It is often used to judge the magnitude of the optical path difference semiquantitatively. A chart for this purpose may be found in Iddings' *Rock Minerals*.

The form of the interference bands depends on the manner in which the optical path difference varies across the surface of the film. The value of θ_2 always varies by at least several degrees over the field of view, so that, in general, $\cos \theta_2$ cannot be regarded as constant even at normal incidence. The importance of the change in the cosine factor depends on the magnitude of $2n_2t$. If this is several thousand wave lengths, one obtains fringes which are due solely to the variation in $\cos \theta_2$, when t is constant. These are called **Haidinger's fringes**, or fringes of constant inclination. They will be discussed further in connection with interferometers. On the other hand, if $2n_2t$ amounts to only a few wave lengths, the variation in the cosine is negligible, and, for incidence around the perpendicular, the cosine may be taken equal to unity. In this case, the dark fringes are seen in an air film when

$$2t = m\lambda$$

or when

$$t = m \frac{\lambda}{2}; \quad m = 0, 1, 2, 3, \dots$$

These fringes are called **fringes of constant thickness**, since the thickness of the film under any one fringe is constant. In a way, the fringes of constant thickness resemble the contour lines on a topographical map, each one of which connects points at a constant elevation of one surface with respect to a second surface, e.g., sea level. There is a constant increment in elevation from one band to its neighbor. Intermediate cases, where $2n_2t$ is neither large nor small, may be discussed in the same manner as given in connection with the Michelson interferometer in Art. 10.9.

10.7 Localization of Fringes

When the film is thin, the fringes are localized in or near the film, as can be seen by noting the intersections of coherent ray pairs reflected at both surfaces, Fig. 10.11(a). For a thick parallel

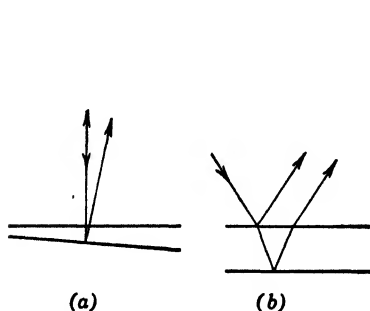


FIG. 10.11. (a) Coherent rays from a wedge-shaped film intersect near the film. (b) Coherent rays from a parallel plate intersect at infinity.

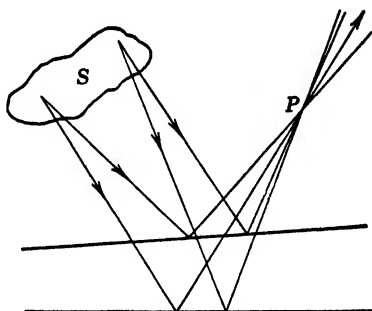


FIG. 10.12. Coherent ray pairs from different points of an extended source arrive at P with different retardations.

plate, on the other hand, the Haidinger's fringes are localized at infinity, since the coherent ray pairs are parallel, Fig. 10.11(b). For intermediate cases, the problem is not solved so simply. Some insight into its nature may be gained by considering the rays in Fig. 10.12. Let S represent the extended light source usually employed, and suppose that one focusses a camera, microscope,

or the eye on some point P . From any point of S one may draw a ray reflected from the upper surface and passing through P , and a coherent ray from the same point of S reflected from the non-parallel lower surface and passing through the same point P . Each ray corresponds to a wave, which is not shown. Besides being easier to trace, the rays indicate more clearly the paths and path differences that determine interference conditions. These rays may be found by an application of Fermat's principle or with the aid of the images of each point in S . The various pairs of rays will possess a certain divergence and difference in path. If the aperture of the observing system is great enough, they will be united at the focus conjugate to P . If one constructs such ray pairs from the various points of S , he may find a considerable variation in the optical path difference to P . If the optical path differences of the united ray pairs vary by less than half a wave length, the interference fringes will be distinct; whereas, if they vary by much more than half a wave length, the distinctness of the fringes is correspondingly reduced.

A restriction in the aperture of the observing system may be necessary in order to limit the variation in optical paths sufficiently to obtain distinct interference bands. The aperture of the eye often provides sufficient limitation. If it does not, a small diaphragm will help. At a particular location of P there will be a maximum distinctness even with a large aperture, and this is where the fringes are said to be localized. This problem will be considered in a quantitative manner in relation with the Michelson interferometer.

10.8 Applications of Fringes of Constant Thickness

Since these fringes are analogous to the contour lines on a topographical map, they may be used to determine the form of a surface of a plate of glass, a prism, or a lens with reference to a standard test plate. For testing surfaces which are approximately plane, one uses an **optical flat**. Such a flat is usually made of Pyrex glass or fused quartz to minimize distortion due to possible unequal temperature. It has one surface which is flat to one fourth, one eighth, or one sixteenth of a wave length of light. The flat surface and the surface to be tested are cleaned, brushed free of dust, and slid into contact. Illumination by monochromatic sodium or mercury light, usually under perpendicular incidence,

reveals bands of constant thickness of air film. These bands are straight and equidistant if the surface being tested is flat. The usual observation of several bands in this case is due to the almost unavoidable wedge-shaped nature of the air film. If the bands

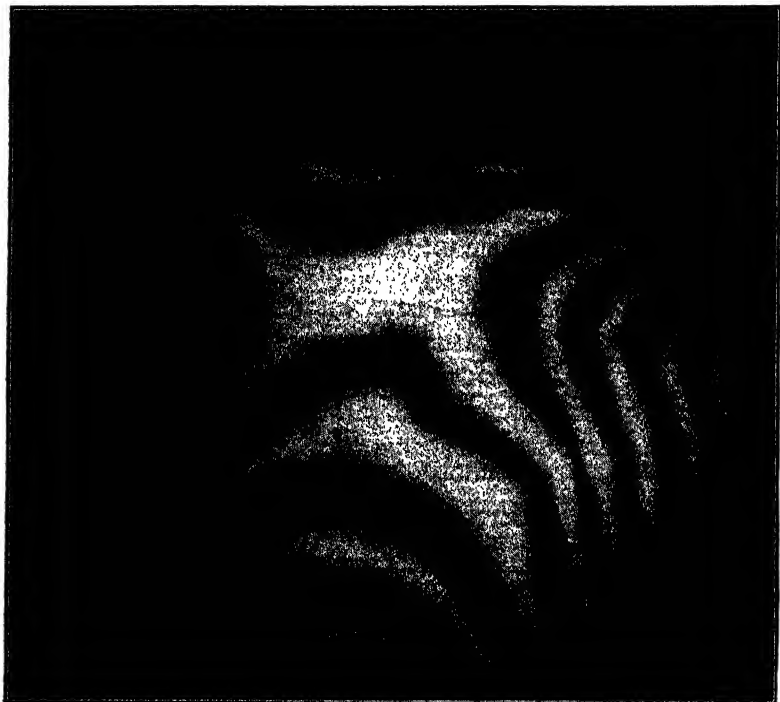


FIG. 10.13. Example of interference bands observed in the air film between a sheet of good plate glass and an optical flat. The plate glass is convex in one principal direction and concave in the other.

have a small curvature, the deviation from flatness across the chord of one band is the same fraction of a half wave length as the ratio of the sagitta of the arc to the interval between bands. If the surface tested deviates appreciably from flatness, irregular bands will be observed over its surface, such as those in Fig. 10.13. Each band is the locus of points at which there is a constant thickness of film, and this thickness changes by half a wave length from any band to the neighboring one. To determine whether the surface slopes downward or upward in a direction normal to the

bands, one may note the direction in which the bands move when a slight pressure is applied, or when the viewing angle is made more oblique. In either case, the bands move toward the place where the air film is thicker, because $2t \cos \theta_2$ remains constant along any one band. Thus a band for a given $2t \cos \theta_2 = m\lambda$ will always be at a place where this condition holds. If t is reduced, while keeping θ_2 constant, the band accordingly moves over to the place where t now has the original value, or where the film *was* thicker. On the other hand, if θ is made larger, $\cos \theta$ becomes smaller, and the band moves over to the place where t is appropriately larger. Thus, if the standard flat is above the sample, the bands move downhill on the surface of the sample. The same is true if the flat is underneath, provided that one imagines the tested surface to be turned so as to face the observer after testing.

If the two surfaces consist of a plane and a sphere, or two spheres of slightly different radius of curvature, one observes rings about the point of contact or nearest approach of the surfaces. These are called **Newton's rings** after their discoverer. Figure 10.14 is a photograph of such rings, using monochromatic mercury light. Let R_1 be the radius of curvature of one of the surfaces, and R_2 the radius of curvature of the other, Fig. 10.15. We wish to find an expression for the thickness of film t at a point P at a distance $r = AP$ from the point of nearest approach of the surfaces or from the axis of symmetry passing through the two centers of curvature. Referring to Fig. 10.15, it is seen that

$$t = AB = AV_2 + V_2V_1 - BV_1$$

The distances AV_2 and BV_1 are sagittas of the two arcs, and are obtained by solving equations like

$$(R_2 - AV_2)^2 + r^2 = R_2^2$$

so that

$$AV_2(2R_2 - AV_2) = r^2$$

When $AV_2 \ll 2R_2$, which is usually the case,

$$AV_2 \doteq \frac{r^2}{2R_2}$$

Similarly

$$BV_1 \doteq \frac{r^2}{2R_1}$$

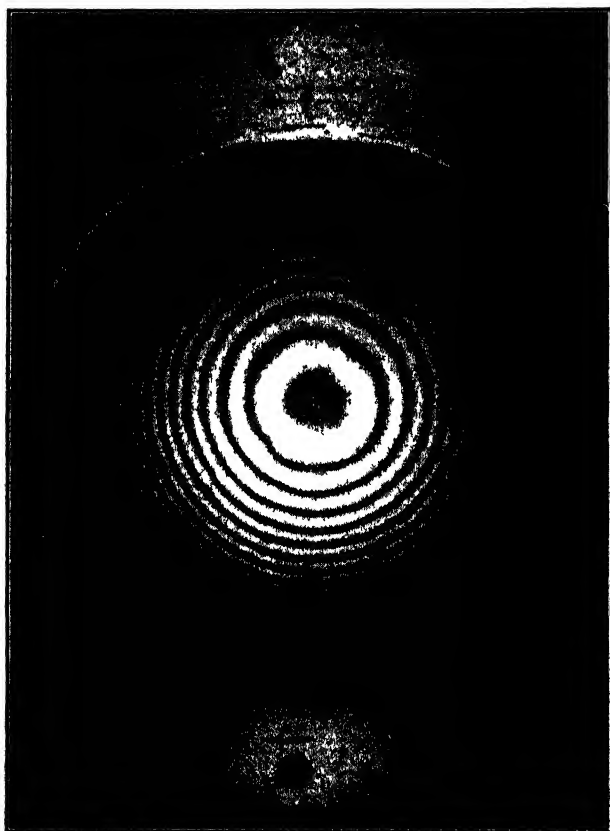


FIG. 10.14. Newton's rings.

Letting e represent the thickness of the film V_1V_2 at the center, it follows that

$$t = \frac{r^2}{2} \left(\frac{1}{R_2} - \frac{1}{R_1} \right) + e$$

Dark rings of radius $r = r_m$ are observed in reflected light when

$$2t \cos \theta = m\lambda$$

or

$$r_m^2 \left(\frac{1}{R_2} - \frac{1}{R_1} \right) \cos \theta + 2e \cos \theta = m\lambda$$

In the special case of $e = 0$, $R_1 = \infty$, and perpendicular incidence, one obtains the usual formula for the radii of Newton's rings;

$$r_m = \sqrt{m\lambda R}$$

The radii of the dark rings are proportional to the square roots of successive integers m , as Newton observed. At the center of the ring system, $m = 0$, a dark spot is observed, which is experimental evidence for a phase reversal. In transmitted light the same formulas hold for the *bright* rings. However, the rings are not nearly so distinct in transmitted light because of the widely different amplitude of the interfering waves. Hence it is best to observe Newton's rings by reflected light.

One may measure the radius of curvature of a lens surface by measuring the diameters of Newton's rings for a known wave length. In practice, it is difficult to be sure that $e = 0$, so it is best to eliminate e by measuring the diameters d_{m_1} and d_{m_2} of two rings. As a matter of fact, e may be quite large, particularly if a concave lens is supported over an optical flat. Eliminating e by using two orders, m_1 and m_2 , one finds that

$$\left(\frac{d_{m_1}^2 - d_{m_2}^2}{4}\right)\left(\frac{1}{R_2} - \frac{1}{R_1}\right) = (m_1 - m_2)\lambda$$

The value of $m_1 - m_2$ is readily found by counting from one ring to the other. For example, if one measures the diameters of the third and thirteenth rings, $m_1 - m_2 = 10$. Then, if one knows the radius of one of the surfaces, often infinite, one may readily compute the radius of curvature of the other.

10.9 Interference with Large Path Difference

In order to obtain distinct interference bands as the path difference becomes large, one must satisfy the following conditions with increasing precision: (a) the light must be monochromatic, and (b) the surfaces must be plane and parallel. A collimator giving incident plane waves provides a larger, more uniform field

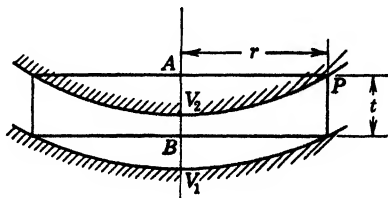


FIG. 10.15. The film thickness t is a function of r .

in which the interference fringes are observed. Various kinds of interferometers have been developed in which the above conditions may readily be satisfied, the best known being the Michelson interferometer. This is quite a versatile instrument and has been used, in various modifications, for many kinds of measurement, some of which are discussed in the following paragraphs.

Plane waves are obtained by placing a source of monochromatic light at the principal focus of a lens L_1 , Fig. 10.16. The waves are

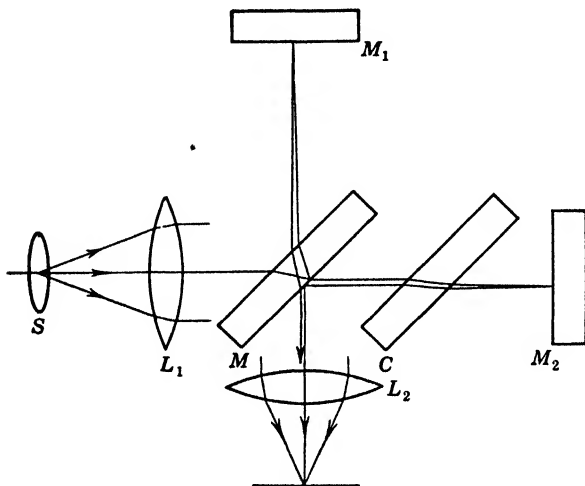


FIG. 10.16. Light path in a Michelson interferometer.

divided at the partially metallized forward surface of M (lower surface in Fig. 10.16) so that coherent waves of approximately equal amplitude travel to the mirrors M_1 and M_2 . Here they are reflected and are finally focussed in the focal plane of the lens L_2 , which may be a telescope objective, a camera lens, or merely the lens system of the observer's eye. The fringes may have various forms, depending on the adjustment of the interferometer, though usually only two forms are used, the straight and the circular. The straight bands are obtained when the difference in path is small, and M_2 is adjusted so that it is not quite at right angles to M_1 . The image of the mirror M_2 in the mirror M , from which the light seems to come, forms a thin wedge with the front surface of M_1 . The fringes are localized near M_1 in this case. If the difference in path is large, the mirrors should be adjusted so that the image of

M_2 in M is strictly parallel to M_1 . In that case, Haidinger's fringes (circles) of constant inclination are observed. These are localized at infinity. The compensator C is made of the same glass and of the same thickness as M , so that it serves to introduce the same amount of material in both the paths. To accomplish this it should be set exactly parallel to M , but it is usually provided with a screw adjustment so that the fringes may be displaced by a small amount when it is rotated. The calibration of the screw is useful in measuring small fringe shifts. The mirror M_1 is mounted

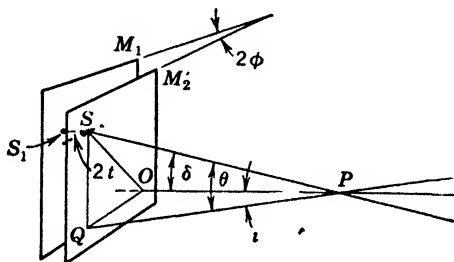


FIG. 10.17. The path difference $2t$ from coherent points in the virtual sources M_1 and M_2' varies with their location.

on accurately straight ways and can be displaced over a range of several centimeters by means of an accurate screw. The number of fringes passing a reference mark in the field of view gives the displacement of the mirror M_1 in half waves of light.

The fringes in the interferometer appear to be formed by light reflected at the surface of M_1 and at the image of M_2 in M . These surfaces generally make a small angle ϕ with each other. As in the theory of interference in thin films, the two reflections introduce a path difference of $2t \cos \delta$ between the two waves, δ being the angle of incidence or reflection, and t being the perpendicular distance between the two surfaces. The general equation for the form and localization of the fringes will now be derived.

If, as is usually true, the source S is of sufficiently large angular extent, the illumination at any external point is independent of its location or form. It is accordingly convenient to imagine the source to lie in the interferometer, being replaced by two coherent sources, one at M_1 and the other in a virtual plane at M_2' , Fig. 10.17. This postulate, due to Michelson, is justified by experience. Thus the light may be considered to come directly from the two

virtual sources M_1 and M_2' without reflection, provided that one doubles the separation t between M_1 and the image of M_2 as well as the angle ϕ between them, as shown in Fig. 10.17. It is easy to see that the path difference $PS_1 - PS$ in the figure is then approximately given by the required expression $2t \cos \delta$.

The phase relations at any point P due to light from coherent points in M_1 and M_2' differ for different pairs of points over the area of M_1 . At the foot of the perpendicular PO of length D the path difference is $2t_0$, where t_0 is the thickness of the equivalent air film between the image of M_2 and the mirror surface M_1 at the point O . Elsewhere the value of $2t$ is given by

$$2t = 2t_0 + OQ \tan 2\phi \doteq 2(t_0 + OQ \tan \phi)$$

the latter approximation being permissible because of the smallness of ϕ . Substituting the value of $OQ = D \tan i$, one obtains for the path difference at P between the light from S and S_1 , the value

$$d = 2t \cos \delta = 2(t_0 + D \tan i \tan \phi) \cos \delta$$

Now

$$\cos \delta = \frac{D}{PS} = \frac{D}{\sqrt{D^2 + OQ^2 + QS^2}} = \frac{1}{\sqrt{1 + \tan^2 i + \tan^2 \theta}}$$

Thus

$$d = 2 \frac{t_0 + D \tan i \tan \phi}{\sqrt{1 + \tan^2 i + \tan^2 \theta}}$$

There will be a maximum distinctness of the fringes at P , i.e., a localization at P , if

$$\frac{\partial d}{\partial \theta} = 0 \quad \text{and} \quad \frac{\partial d}{\partial i} = 0$$

From the first condition, one finds by differentiation that

$$\frac{\partial d}{\partial \theta} = - \frac{2(t_0 + D \tan i \tan \phi)}{2(1 + \tan^2 i + \tan^2 \theta)^{3/2}} \cdot 2 \tan \theta \sec^2 \theta = 0$$

Hence

$$\theta = 0$$

Moreover, from the second condition,

$$\frac{2D \sec^2 i \tan \phi}{(1 + \tan^2 i + \tan^2 \theta)^{1/2}} - \frac{t_0 + D \tan i \tan \phi}{(1 + \tan^2 i + \tan^2 \theta)^{3/2}} 2 \tan i \sec^2 i = 0$$

or

$$\frac{2 \sec^2 i}{(1 + \tan^2 i + \tan^2 \theta)^{3/2}} (D \tan \phi + D \tan \phi \tan^2 \theta - t_0 \tan i) = 0$$

and

$$D \tan \phi (1 + \tan^2 \theta) = t_0 \tan i$$

Since $\theta = 0$, it follows that

$$D \tan \phi = t_0 \tan i$$

Hence

$$D = \frac{t_0 \tan i}{\tan \phi}$$

giving the distance from M_1 at which the fringes are localized. Two special cases are frequently encountered, namely,

1. if $t_0 \neq 0$ and $\phi = 0$, then $D = \infty$, the fringes being localized infinitely far away, while
2. if $\phi \neq 0$, but $t_0 = 0$, then $D = 0$, and the fringes are at the film.

For any point of observation at a distance D from M_1 , the locus of points of constant d on M_1 relative to coherent points on M_2' will be given by an equation of the same form as derived for d in the preceding paragraph. The form of the fringes on M_1 will be given by the coordinates x, y , when d is any constant, where

$$x = D \tan i$$

$$y = D \tan \theta$$

In the following discussion D is the distance from the first nodal point of the observing lens L_2 to M_1 , or one may consider it to be the distance from the second nodal point to the plane on which the fringes are projected. Then

$$d^2(1 + \tan^2 i + \tan^2 \theta) = 4(t_0 + D \tan i \tan \phi)^2$$

or

$$d^2 \left(1 + \frac{x^2}{D^2} + \frac{y^2}{D^2} \right) = 4(t_0^2 + x^2 \tan^2 \phi + 2t_0 x \tan \phi)$$

or

$$x^2(d^2 - 4D^2 \tan^2 \phi) + y^2 d^2 - 8xt_0 D^2 \tan \phi + d^2 D^2 - 4D^2 t_0^2 = 0$$

which is the equation of a conic section. Thus the fringes may be straight lines, circles, parabolas, ellipses, or hyperbolas, depending on the values of d and ϕ : Two cases are of especial interest, namely, (1) $t_0 = 0$, or at most a few wave lengths, and (2) $\phi = 0$.

In the first case, it is advantageous to start with the original equation for d . Since $\tan i = x/D$ and $\tan \theta = y/D$ are small compared with unity, one may reduce this equation to the simple form

$$d = 2(t_0 + x \tan \phi) \cong 2x \tan \phi$$

The dark bands occur where $d = m\lambda$, and therefore at

$$x = m \left(\frac{\lambda}{2 \tan \phi} \right)$$

This is the equation of a family of straight lines, the separation of which increases as ϕ becomes smaller.

In the second case, $\phi = 0$, the quadratic formula reduces, without approximations, to

$$x^2 d^2 + y^2 d^2 + D^2(d^2 - 4t_0^2) = 0$$

or

$$x^2 + y^2 = D^2 \left(\frac{4t_0^2 - d^2}{d^2} \right) = \text{constant}$$

which is the equation of a family of circles, the dark circles being defined by $d = m\lambda$. The radii of the circles increase as d decreases. As d approaches zero, the entire field should be dark, but this rarely happens, since variations in thickness of the metallic coatings, deviations from planeness of the glass plates, and non-homogeneity in the diagonal plates cause the path difference to vary over the field. Thus streaks of varying intensity are observed over the field as d approaches zero.

10.10 The Michelson-Morley Experiment

The Michelson interferometer has been used in many important investigations, the most striking of which is the Michelson-Morley experiment, in which an unsuccessful attempt was made to measure the speed of the ether "wind" past the moving earth. If one arm of the interferometer is in the direction of the earth's motion relative to the "ether" and the other is at right angles to

this motion, the relative retardation in units of time is found by simple kinematics to be

$$\Delta t = L \left(\frac{1}{c-v} + \frac{1}{c+v} - \frac{2}{\sqrt{c^2-v^2}} \right)$$

where L is the length of each interferometer arm, v is the velocity of the interferometer, and c is the velocity of light. The relative retardation in units of length is then

$$d = c \Delta t = 2L \left(\frac{1}{1 - \frac{v^2}{c^2}} - \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \right)$$

Since v/c is small, one may expand the fractions in the parentheses by the binomial theorem, retaining only the first two terms, thus finding that

$$d \cong L \frac{v^2}{c^2}$$

By turning the interferometer through 90° the arms are interchanged, and the fringe shift expected is

$$\frac{2d}{\lambda} = \frac{2L}{\lambda} \frac{v^2}{c^2}$$

At some season of the year, the motion of the earth with respect to the sun might be compensated by the motion of the solar system, so it is essential to make observations at various times of the year. Since $v^2/c^2 = 10^{-8}$ for the known orbital velocity of the earth, the order of magnitude of the expected fringe shift is

$$\frac{2L}{\lambda} \times 10^{-8}$$

In order to measure such a small shift, it is necessary to make $2L$ equal to about 10^8 wave lengths. This was accomplished by using a large interferometer on a float supported by mercury, and by reflecting the light back and forth four times to make the equivalent value of L equal to about 3403 cm, so that

$$\frac{2L}{\lambda} \cong 1.12 \times 10^8$$

which would give an easily measured fringe shift. Although the observed shift was never zero, and varied in a curious manner with the season of the year, it was always less than 5 per cent of that to be expected.¹ According to the theory of relativity, there should be no observable effect at all. This may be considered to be practically the case. On the other hand, the actually observed fringe shift, if real, is of considerable interest.

The presumably negative result of this experiment served as an impetus to the development of the theory of relativity by Einstein. For an interesting account of this theory, the reader is referred to Lieber and Lieber's *Relativity*.

10.11 Visibility of the Fringes

If the incident light consists of two or more wave lengths, the visibility V of the interference fringes will change periodically with the plate separation t , as shown in Fig. 10.18 for the simple

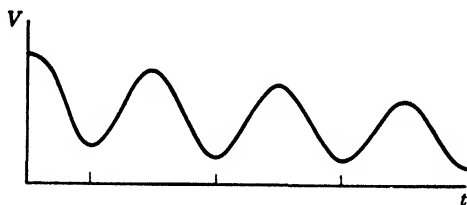


FIG. 10.18. The visibility of interference fringes with a source emitting two wave lengths varies periodically with the path difference.

case of two spectral lines. The visibility of the interference pattern is defined by the equation

$$V = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}}$$

where I_{\max} is the intensity of the interference maxima, and I_{\min} is the intensity of the minima. As a simple illustration, assume two wave lengths λ_1 and λ_2 of approximately equal amplitude a . The interference pattern of λ_1 can be described by

$$I_1 = 2a^2(1 + \cos \Delta_1) \quad \text{where} \quad \Delta_1 = \frac{2\pi}{\lambda_1} t$$

¹ D. C. MILLER, *Reviews of Modern Physics*, 5, 203-242 (1933).

and that of λ_2 by a similar equation

$$I_2 = 2a^2(1 + \cos \Delta_2) \quad \text{where} \quad \Delta_2 = \frac{2\pi}{\lambda_2} 2t$$

The superposed intensity variation is given by $I = I_1 + I_2$, since the two waves of different lengths are not coherent. Thus

$$I = 2a^2[2 + (\cos \Delta_1 + \cos \Delta_2)]$$

or

$$I = 4a^2 \left(1 + \cos \frac{\Delta_1 + \Delta_2}{2} \cos \frac{\Delta_1 - \Delta_2}{2} \right)$$

Since the wave lengths are nearly equal, we will write Δ for the average phase difference, so that

$$I = 4a^2 \left(1 + \cos \Delta \cos \frac{\Delta_1 - \Delta_2}{2} \right)$$

The quantity Δ varies much more rapidly than $(\Delta_1 - \Delta_2)/2$, giving an interference fringe pattern which is modulated by the second factor. The fringes vanish when this second cosine factor is zero, making $I = 0$, since $I_{\max} = I_{\min}$, when

$$\cos \frac{\Delta_1 - \Delta_2}{2} = 0 \quad \text{or} \quad \frac{\Delta_1 - \Delta_2}{2} = \left(m + \frac{1}{2} \right) \pi$$

Substitution of the values of Δ_1 and Δ_2 leads to

$$2t \left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2} \right) = m + \frac{1}{2}$$

as the condition for zero visibility of the fringes, or

$$2t = \left(m + \frac{1}{2} \right) \left(\frac{\lambda_1 \lambda_2}{\lambda_2 - \lambda_1} \right)$$

By observing the periodic disappearance of the fringes at intervals $2t_i$, one may determine the difference between the wave numbers, since

$$\frac{1}{\lambda_1} - \frac{1}{\lambda_2} = \frac{1}{2t_i}$$

If the wave lengths are close enough together, one may write for their separation

$$\Delta\lambda \doteq \frac{\lambda^2}{2t_i}$$

If the amplitudes of the two waves are not exactly equal, the visibility of the fringes will not pass through zero, but through minima at the same equal intervals.

The visibility curves for more general cases have been discussed in detail by Michelson, who analyzed such curves to find the structure of spectral lines in the days before one could obtain sufficient resolving power to observe them directly.

10.12 Widths of Spectral Lines

If spectral lines were perfectly sharp, that is to say, if the light were perfectly monochromatic, the visibility of the fringes in Michelson's interferometer would remain constant as the difference in path is increased. For lines of finite width, the visibility drops off, as Michelson showed.² The variation of visibility with path difference is found to depend on the form and width of the spectral line.

Let $\psi(x)$ represent the distribution function of intensity over a spectral line whose center, at wave number $\bar{\nu}$, is the origin from which x is measured. Then $1/\lambda = \bar{\nu} + x$, and the phase difference for any element at x is

$$\Delta = \frac{2\pi D}{\lambda} = 2\pi D(\bar{\nu} + x)$$

where D is written for the path difference to avoid confusion with the differential operator d . Assuming that the half-silvered mirror of the interferometer divides the amplitude into two equal parts, the intensity in the interference pattern due to the portion of the spectral line between x and $x + dx$ is

$$dI = 2a^2(1 + \cos \Delta)$$

which becomes in this case

$$dI = \psi(x) (1 + \cos \Delta) dx$$

² MICHELSON, *Studies in Optics*, University of Chicago Press, 1927, Chapter

Substituting the value of Δ , one finds that

$$dI = \psi(x) dx (1 + \cos 2\pi D\bar{v} \cos 2\pi Dx - \sin 2\pi D\bar{v} \sin 2\pi Dx)$$

and, consequently,

$$I = \int_{-\infty}^{+\infty} \psi(x) dx + \cos 2\pi D\bar{v} \int_{-\infty}^{+\infty} \psi(x) \cos 2\pi Dx dx \\ - \sin 2\pi D\bar{v} \int_{-\infty}^{+\infty} \psi(x) \sin 2\pi Dx dx$$

This may be written in the form

$$I = P + C \cos 2\pi D\bar{v} - S \sin 2\pi D\bar{v}$$

by obvious substitution of P , C , and S for the three integrals. Since $\bar{v} \gg x$ over the line, one may consider C and S to be constant while finding the conditions for maximum and minimum intensity in the interference pattern by setting

$$\frac{\partial I}{\partial D} = 0$$

Thus one finds that

$$C \sin 2\pi D\bar{v} + S \cos 2\pi D\bar{v} = 0$$

or

$$\tan 2\pi D\bar{v} = -\frac{S}{C}$$

Therefore

$$\sin 2\pi D\bar{v} = -\frac{S}{\pm \sqrt{S^2 + C^2}}$$

and

$$\cos 2\pi D\bar{v} = \frac{C}{\pm \sqrt{S^2 + C^2}}$$

so that

$$I_m = P \pm \sqrt{C^2 + S^2}$$

From this it follows that

$$I_{\max} = P + \sqrt{C^2 + S^2}$$

and

$$I_{\min} = P - \sqrt{C^2 + S^2}$$

From the definition of visibility V , one finds that

$$V = \frac{\sqrt{C^2 + S^2}}{P}$$

If the line is symmetrical about \bar{v} , the integral S is zero and the expression for V reduces to

$$V = \frac{C}{P} = \frac{\int_{-\infty}^{+\infty} \psi(x) \cos 2\pi D x \, dx}{\int_{-\infty}^{+\infty} \psi(x) \, dx}$$

from which V may be obtained for any given $\psi(x)$. It is found that V is a function of D , the visibility dropping off, with or without maxima, as the difference in path is increased.

In particular, if $\psi(x)$ is of the nature of a probability curve

$$\psi(x) = e^{-x^2/w^2}$$

the visibility function is also found to have the form of a probability curve

$$V = e^{-(\pi D w)^2}$$

A broader $\psi(x)$ leads to a more rapid decline in V . From these two equations one readily finds that the half width of the line $x_{1/2}$ at half its maximum intensity is related to the path difference $D_{1/2}$ at which the visibility drops to half value, by the equation

$$x_{1/2} = \frac{\log_e 2}{\pi D_{1/2}} = \frac{0.22}{D_{1/2}}$$

Note that the full width of the line at half maximum intensity is twice $x_{1/2}$ and that the result is in reciprocal centimeters or wave number units.

Quantum states have a natural width which is inversely proportional to their lifetime. This depends on the transition probabilities to all states which may combine with a given state. The half width of a spectrum line is equal to the sum of the two term widths corresponding to the initial and final states. If a term combines with a continuous band of energy levels, all lines involving this term will be abnormally broad. In most cases, however, the

natural width is so much smaller than that due to other causes that it may be disregarded.

Thus in practice the usual causes of line breadth are collision broadening and the Doppler-Fizeau effect, Art. 17.1. The effect of collisions is to shorten the wave trains and consequently broaden the frequency distribution, as one may infer from Fourier's integral. If l is the mean free path, T the absolute temperature, and M the molecular weight, the effect of collisions is to make

$$x_{1/2} = \frac{0.11 \times 10^{-6}}{l} \sqrt{\frac{T}{M}} \text{ cm}^{-1}$$

The corresponding expression for the broadening due to the Doppler-Fizeau effect is

$$x_{1/2}' = 0.39 \times 10^{-6} \bar{\nu} \sqrt{\frac{T}{M}} \text{ cm}^{-1}$$

where $\bar{\nu}$ is the wave number of the center of the line. It is evident that both effects increase in the same way with temperature and decrease in the same way as the mass of the ion increases. One may verify these conclusions with a Michelson interferometer. The relative importance of these two causes of line broadening depends on the mean free path and therefore on the pressure. Their ratio is

$$\frac{x_{1/2}}{x_{1/2}'} = \frac{1}{3.5\bar{\nu}l} = \frac{\lambda}{3.5l}$$

At 760 mm pressure, at which $l = 0.04 \mu$, it is evident that collision broadening will have the larger effect on the width of spectral lines, whereas in discharge tubes at low pressure the Doppler broadening is larger.

It should not be overlooked that lines may also appear broad because of unresolved hyperfine structure or a Stark effect due to electric fields. Hyperfine structure may be due either to the presence of isotopes in sufficient abundance to emit observable lines, or to the various modes of coupling of nuclear spin with the resultant angular momentum of the rest of the atom. The latter leads to a multiplicity of lines in the case of even a single isotope.³

³ HERZBERG, *Atomic Structure and Atomic Spectra*, Prentice-Hall, 1937, Chapter 5.

Michelson's studies showed that the red line of cadmium is remarkably free from disturbing satellites and is very narrow. For this reason it was selected for the measurement of the standard meter in terms of light waves. The reciprocal of this result has been adopted as the primary wave length in spectroscopy. Its value is 6438.4696 international angstrom units for the wave length of the cadmium red line in dry air at normal pressure and a temperature of 15°C. Recently one of the isotopes of mercury, of mass number 198, has been produced in sufficient quantity by neutron bombardment of gold, so that it may be used by itself in a mercury discharge tube. The green line radiated by this isotope is more monochromatic than any hitherto obtained. Because the operation of this source is more convenient than that of the cadmium arc, it promises to displace the latter as a primary standard.⁴

10.13 Jamin Interferometer

Another useful type of interferometer is that of Jamin, of which modifications have been devised by Mach and by Zehnder. One

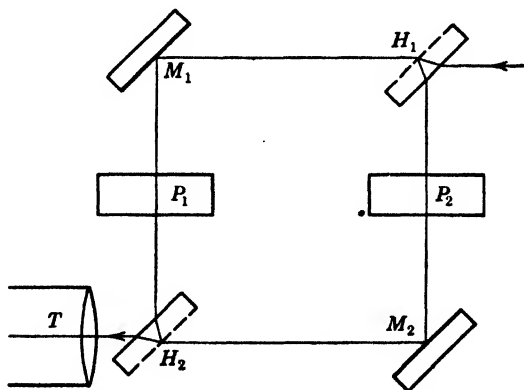


FIG. 10.19. A form of Jamin interferometer due to Mach and Zehnder. M_1 , M_2 , H_1 , and H_2 are mirrors, H_1 and H_2 being half-silvered. P_1 and P_2 are plates of material being compared. Interference bands are observed in the telescope T .

typical form is shown in Fig. 10.19. The incident light is separated into two beams of approximately equal intensity at the half-silvered or, better, aluminized surface H_1 . After travelling along

⁴ MEGGERS, *Journal of the Optical Society of America*, **38**, 7-14 (1948).

the paths as indicated, the light is reunited in the focal plane of the telescope T . If one of the mirrors is inclined slightly, the waves arriving at the telescope give rise to fringes of the form explained in the discussion of interference. This interferometer has the advantage that light passes just once through any one of its arms. It was used by Mach to study the flow of air past flying bullets and is now being widely employed to study air flow past airfoils and other obstacles at supersonic speeds. Because $n - 1$ for air is proportional to its density, one may readily translate fringe shifts into density changes. This form of interferometer has also been used to find the change in index of refraction of transparent materials when subjected to mechanical stresses.

If two identical plates P_1 and P_2 of some material of index n_0 and thickness t are placed in the two beams as shown in Fig. 10.19, there will be no change in optical paths and the interference bands will not be displaced. If now the index of refraction of one of the plates is changed by some external agency such as an applied stress, the interference bands will be displaced by an amount which depends on the phase difference Δ produced by this change in refractive index, where

$$\Delta = \frac{2\pi}{\lambda} (n - n_0)t$$

To measure this, it is necessary to calibrate the interferometer by measuring the interval s_0 between successive fringes by means of cross-hairs and a micrometer. This interval corresponds to a change in phase of just 2π . Any shift in the fringes s is to the fringe interval s_0 as the phase shift causing it is to 2π , or

$$\Delta = \frac{s}{s_0} 2\pi$$

If polarized light is used to illuminate the interferometer, it should vibrate in the plane of incidence or at right angles to it in order that it remain plane polarized. Using such light, one may show that the change in phase is different for light vibrating in the two principal vibration directions of the plate.

10.14 The Fabry-Perot Interferometer

The Fabry-Perot interferometer consists essentially of two plane reflecting flats, the separation of which is controlled by an ac-

curate screw. Lens L_1 , Fig. 10.20, collimates the light from a broad source S in its principal focal plane. Multiple reflections take place between the metallized surfaces of M_1 and M_2 . These should be of as high a reflecting power as is consistent with sufficient transmission to observe the interference bands. The bands are formed by the lens L_2 in its focal plane and consist of circular "fringes of constant inclination."

At any point such as P there are united the multiply reflected waves which have a common inclination ϕ . Intensity maxima

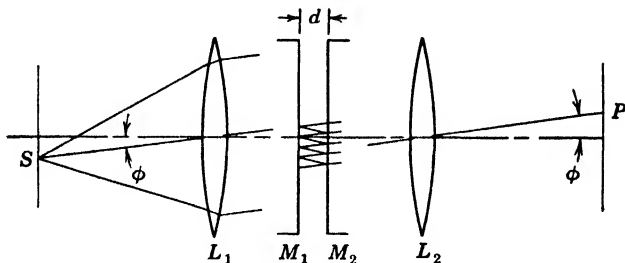


FIG. 10.20. Light path in a Fabry-Perot interferometer.

occur when the difference in path for any two successive reflections is an integral multiple of a wave length, namely, when

$$2d \cos \phi = p\lambda$$

where p is an integer. These circular intensity maxima are quite sharp, so that very small wave length differences may be resolved if the plate separation d is great enough. This makes the interferometer suitable for the study of hyperfine structure of spectral lines, which provides data pertaining to the nuclear spin of the atom.

The amplitude at P is found by adding the waves represented by

$$s_1 = Q \sin \omega t$$

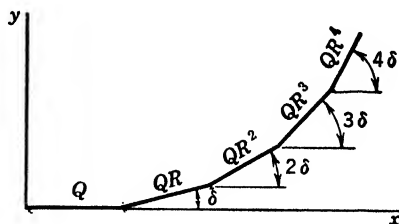
$$s_2 = QR \sin (\omega t - \delta)$$

$$s_3 = QR^2 \sin (\omega t - 2\delta), \text{ etc.}$$

where R is the intensity reflection coefficient of the mirrors M_1 and M_2 , which are assumed to reflect equally. Since there are two added reflections for each successive wave, the amplitude is re-

duced by the square (product if different) of the amplitude reflection coefficient at the two reflecting surfaces, the square being equal to the intensity reflection coefficient R . The transmission coefficient for the second surface is included in the factor Q . The phase changes by increments of

$$\delta = \frac{2\pi}{\lambda} 2d \cos \phi$$



for each successive transmitted wave. FIG. 10.21. Vibration polygon for the multiple waves from a Fabry-Perot interferometer.

The amplitude polygon for the summation of the multiple waves starts off as shown in Fig. 10.21. The intensity is proportional to the square of the closing side, which is the sum of the squares of its x and y components. These are given by

$$\Sigma x = Q(1 + R \cos \delta + R^2 \cos 2\delta + \dots)$$

and

$$\Sigma y = Q(0 + R \sin \delta + R^2 \sin 2\delta + \dots)$$

Hence the intensity I is found by the operation

$$I \propto (\Sigma x)^2 + (\Sigma y)^2 = (\Sigma x + i \Sigma y)(\Sigma x - i \Sigma y)$$

where $i = \sqrt{-1}$, so that

$$\Sigma x + i \Sigma y = Q(1 + Re^{i\delta} + R^2 e^{2i\delta} + \dots) = \frac{Q}{1 - Re^{i\delta}}$$

and

$$\Sigma x - i \Sigma y = Q(1 + Re^{-i\delta} + R^2 e^{-2i\delta} + \dots) = \frac{Q}{1 - Re^{-i\delta}}$$

Thus it is found that

$$\begin{aligned} I &= \frac{kQ^2}{1 + R^2 - 2R \cos \delta} = \frac{kQ^2}{(1 - R)^2 + 4R \sin^2 \frac{\delta}{2}} \\ &= \frac{I_m}{1 + \frac{4R}{(1 - R)^2} \sin^2 \frac{\delta}{2}} \end{aligned}$$

where

$$I_m = \frac{kQ^2}{(1 - R)^2}$$

is the maximum intensity which occurs when $\delta/2 = p\pi$ or

$$\frac{\pi}{\lambda} 2d \cos \phi = p\pi$$

or

$$2d \cos \phi = p\lambda$$

The maxima are quite sharp if R is large. For example, if $R = 0.8$, then

$$\frac{4R}{(1 - R)^2} = \frac{3.2}{(0.2)^2} = 80$$

and

$$I = \frac{I_m}{1 + 80 \sin^2 \delta/2}$$

The intensity I will accordingly drop to $\frac{1}{2}I_m$ when $\sin^2 \delta/2 = 1/80$ or when

$$\delta = 2 \sin^{-1} (80)^{-1/2} = 2(p\pi \pm 0.112) = p \cdot 2\pi \pm 0.224$$

The quantity 0.224 is a small fraction of the total change in δ of 2π radians between orders. The maxima become still sharper with increasing R . Note that the intensity does not fall to zero between maxima, but to a minimum of

$$I_{\min} = \frac{I_{\max}}{1 + \frac{4R}{(1 - R)^2}}$$

which is equal to one eighty-first of I_m if $R = 0.8$.

The resolving power of any spectroscope is defined by

$$R = \frac{\lambda}{\Delta\lambda_m}$$

where $\Delta\lambda_m$ is the smallest change in the wave length λ that can be resolved. In the case of an interferometer, when $\cos \phi \doteq 1$,

$$P\lambda = 2d$$

defines a variable order P . The change in order corresponding to a change in λ is obtained from the differential formula

$$P \Delta\lambda + \lambda \Delta P = 0$$

so that

$$\frac{\lambda}{\Delta\lambda_m} = \frac{P}{\Delta P_m}$$

The smallest change in P that may be resolved is proportional to the change in phase difference $\Delta\delta_m$ that can be detected. Thus

$$\frac{\Delta P_m}{1} = \frac{\Delta\delta_m}{2\pi}$$

as may be seen by referring to Fig. 10.22. From a study of the forms of the maxima, it is concluded that two neighboring maxima cross at $I = 0.405I_{\max}$ when they are just resolvable, so that the problem reduces to that of finding the change in δ for which the intensity becomes $0.405I_{\max}$. Since all maxima have the same

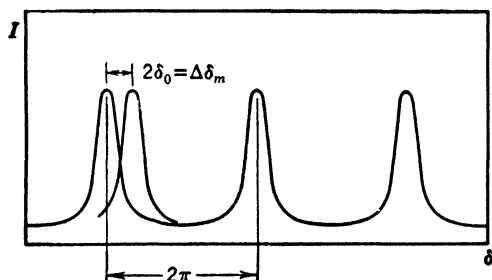


FIG. 10.22. Form of interference maxima in a Fabry-Perot interferometer. Applied to discussion of resolving power.

form, the change in δ is equal to $2\delta_0$ for the zero order peak, where δ_0 is the phase difference for $I/I_m = 0.405$ in this peak. Thus

$$1 + \frac{4R}{(1-R)^2} \sin^2 \frac{\delta_0}{2} = \frac{1}{0.405}$$

$$\sin^2 \frac{\delta_0}{2} = 0.368 \frac{(1-R)^2}{R}$$

$$\delta_0 = 2 \sin^{-1} \left[\frac{0.368(1-R)^2}{R} \right]^{1/2}$$

Hence the expression for the resolving power of a Fabry-Perot interferometer is

$$R = \frac{2\pi P}{2\delta_0} = \frac{\pi P}{2 \sin^{-1} \left[\frac{0.368(1 - R)^2}{R} \right]^{1/2}}$$

If the spacing d is 10 mm and the wave length is 5000 Å, then for $R = 80$ per cent, $R = 465,400$; and for $R = 90$ per cent, $R = 1,047,200$.

The advantage of a high reflectance is apparent. The best material to use in the visible region is silver applied by evaporation in a vacuum. In the ultraviolet, silver is quite transparent near 3100 Å, so that another metal such as aluminum is frequently used. Because silver tarnishes rapidly on exposure to the atmosphere, an aluminum mirror is often more practical in the visible region as well, if the resolving power it yields, because of its lower reflectance, is high enough.

At the center of the ring system, one may write $2d = P\lambda$, where P is not an integer in general. For the smallest ring, P exceeds p_1 by some fraction ϵ_1 , called the **fractional excess**. In fact, for any ring

$$P = p + \epsilon$$

The ratio of two wave lengths may be determined by finding the ratio of the P 's by

$$\frac{\lambda_1}{\lambda_2} = \frac{P_2}{P_1}$$

Since the angle ϕ is quite small for the smaller rings, one may write

$$2d \cos \phi \doteq 2d \left(1 - \frac{\phi^2}{2} \right) \doteq p\lambda$$

or

$$P \left(1 - \frac{\phi^2}{2} \right) = p$$

so that

$$P = \frac{p}{1 - \frac{\phi^2}{2}}$$

Thus

$$\frac{\lambda_1}{\lambda_2} = \frac{p_2}{p_1} \frac{\left(1 - \frac{\phi_1^2}{2}\right)}{\left(1 - \frac{\phi_2^2}{2}\right)}$$

The values of ϕ for the smallest rings due to the two wave lengths are determined by measuring the diameters D of the rings in the focal plane of the observing lens, whose focal length f should be accurately known. Then

$$\phi = \frac{D}{2f}$$

so that

$$\frac{\lambda_1}{\lambda_2} = \frac{p_2}{p_1} \left(1 - \frac{D_1^2}{8f^2} + \frac{D_2^2}{8f^2}\right)$$

The wave lengths are often known accurately enough so that p_2/p_1 may be computed with sufficient precision. Or, one may find the proper values of p_1 and p_2 by using the moderately accurate ratio λ_1/λ_2 in computing

$$P_2 = \frac{\lambda_1}{\lambda_2} (p_1 + \epsilon_1)$$

into which one substitutes the probable value of p_1 and the observed fractional excess ϵ_1 . If P_2 does not come out with the observed fractional excess, one changes the value of p_1 by integral steps until the computed P_2 has the observed fractional excess ϵ_2 . It is advisable to work with several lines simultaneously, computing their values of P_3, P_4, P_5 , etc., in order to avoid a fortuitous agreement of computed and observed fractional excesses which may sometimes be found with a single pair of lines. The values found for p_1, p_2 , etc., which give the best agreement are then used in the above formula for the ratio of the wave lengths in terms of the diameters of the rings or the observed fractional excesses. For further experimental details and for a discussion of other forms of interference spectroscopes, such as the Lummer-Gehrcke plate and the Michelson echelon grating, the reader is referred to Tolansky's *High Resolution Spectroscopy*, Williams' *Applications of Interferometry*, Bruhat's *Cours d'Optique*, or Jenkins and White's *Fundamentals of Physical Optics*.

PROBLEMS

1. What is the amplitude of the wave resulting from the superposition of two waves vibrating in the same plane if their amplitudes are respectively 10 and 5 statvolts per centimeter, and the phase difference is $\pi/4$?

2. Two plane surfaces of glass in contact along one edge are separated at the opposite edge by a thin fiber. If twenty interference fringes are observed between these edges when perpendicularly reflected sodium light is used, what is the thickness of the fiber?

3. An interference device is used to observe the expansion of a metal bar. If 150 bands move past the cross-hairs in the observing telescope during the run, what is the change in length of the bar, the wave length of the light used being 5461 Å?

4. A standard Johansson gauge block and a slightly taller unknown block are wrung onto a base plate adjacent to each other. A plane test plate rests on their upper surfaces. If eight dark interference fringes are seen above the standard gauge in perpendicularly reflected sodium light, how much taller is the adjacent block?

5. Interference rings are observed between a convex surface in contact with a plane surface illuminated normally by sodium light. If the diameter of the twentieth dark ring in reflected light is 6 mm, what is the radius of curvature of the convex surface?

6. In a Newton's rings experiment, a concave spherical surface is supported above an optical flat. The diameter of the fifth dark Newton's ring in perpendicularly reflected sodium light is 2.25 mm, and that of the fifteenth dark ring is 4.25 mm. What is the radius of curvature of the concave surface?

7. (a) Write the conditions for maximum and for minimum intensity of reflection by two parallel plane surfaces when $n_1 < n_2 > n_3$, and also when $n_1 < n_2 < n_3$. (b) Where are the fringes localized in the case of a very thin wedge and also in the case of a parallel plate? (c) Explain the colors of interference fringes in white light.

8. White light is reflected perpendicularly at the surfaces of an air film $2\ \mu$ thick between glass plates. Which wave lengths in the visible spectrum will be of minimum intensity?

9. What should be the index of refraction and the thickness of the thinnest film on glass of index 1.65 such that it will not reflect light 5500 Å in wave length? What would be the color of this film in white light? Why would a thicker film be undesirable as a low-reflectance coating on glass?

10. What would be the fringe displacement for a wave length of 5461 Å when one of two identical 50-cm-long tubes containing H_2 ($n = 1.000139$) in the respective arms of a Rayleigh gas interferometer is evacuated?

11. White light is transmitted through a Fabry-Perot etalon having a plate separation of $6\ \mu$. At what wave lengths in the spectrum between 5790 Å and 4916 Å will the intensity maxima lie?

12. If twenty-one interference fringes due to selective constructive interference in a Fabry-Perot etalon are observed in a channelled spectrum between 5790 Å and 4358 Å, what is the separation of the plates?

13. Prove the formula on page 124 for the increase in optical path produced by rotating a compensator of thickness t and index n through an angle i , starting from normal incidence.

14. Determine the form of the visibility curve in Michelson's interferometer if $\psi(x) = 1$ for $-w/2 < x < +w/2$ and $\psi(x) = 0$ everywhere else.

15. (a) Determine the form of the visibility curve in Michelson's interferometer if $\psi(x) = Ce^{-x^2/w^2}$. (b) If the visibility drops to half its maximum value in a path difference of $D = 5$ cm, what is the full width of the spectrum line at half maximum intensity? Express the result in angstrom units at a wave length of 5000 \AA .

16. What is the resolving power of a Fabry-Perot interferometer when the plate separation is 12 mm and the reflecting power of both plates is 85 per cent? Assume a wave length of 5000 \AA .

Chapter 11. Diffraction

11.1 Diffraction Phenomena

It has been pointed out that the principle of rectilinear propagation is not strictly true, but only appears to hold if the light is not restricted by small apertures. This chapter will be concerned with these deviations of light from a rectilinear path. Such phenomena are called **diffraction effects**. They are characteristic of all kinds of wave motion. Thus it is common experience that water waves, sound waves, and radio waves bend around obstacles so that no clear-cut "shadows" are observed. Light, on the other hand, seems to behave differently, but we shall see that the small amount of diffraction is due to the wave length usually being small compared with the sizes of obstacles and apertures commonly encountered. When the wave length is not of negligible magnitude as compared with the obstacles and apertures, the same diffraction effects are observed as in the case of sound or other waves where similar conditions hold. Excellent photographs of optical diffraction phenomena have been published by Arkadiev in the *Physikalische Zeitschrift*, 45, 832.

Diffraction of light was discovered by the Jesuit Father Francesco Grimaldi and described in his book, *Physico-mathesis de Lumine, Coloribus, et Iride*, which was published by a friend in 1665, two years after Grimaldi's death. Newton discusses diffraction in his *Opticks* and tries to explain it in terms of forces between matter and the light corpuscles travelling near it. He seems to have been unaware of the similarity of optical diffraction to that of sound and other waves. To Fresnel goes the credit for developing, in 1814-1816, the first satisfactory theory of diffraction by combining the principle of interference with Huygens wavelets. It is now possible to present a more elegant, and difficult, mathematical treatment of diffraction by the use of Kirchhoff's solution of the wave equation. But, in spite of the mathematical elegance of Kirchhoff's theory, its fundamental assumptions have been often criticized, as have those of Fresnel. On the other hand, in regard to computed intensity distributions, either

theory leads to useful results which are in general agreement with experiment. Hence the simpler ideas of Fresnel will be used in developing the form of these distributions, their criticism being left for discussion in Art. 11.8.

Diffraction phenomena are classified into two categories. When the waves are spherical, we have what are called **Fresnel diffraction phenomena**. Experimentally, they are the simplest to observe, since no lenses, just a narrow source, diffracting aperture, and screen are required. However, the fact that the waves are spherical makes the mathematical theory more complicated than that of the **Fraunhofer diffraction phenomena** which are observed when the waves are plane. These phenomena are observed in the laboratory when an aperture is mounted on the table of an adjusted spectrometer or when some equivalent arrangement is used. The waves incident on the aperture are plane waves, and the diffracted waves are again focussed in the focal plane of an observing telescope or camera lens where each point corresponds to a conjugate plane in the object space. Fraunhofer diffraction has more important applications than the Fresnel type, being the foundation of the theory of resolving power and of diffraction gratings. We shall therefore reverse the usual order of presentation and discuss the mathematically simpler Fraunhofer diffraction first.

11.2 Single Slit in Parallel Light

Let AB , Fig. 11.1, represent a slit mounted on the table of a spectrometer. In practice, the angles of incidence i and of dif-

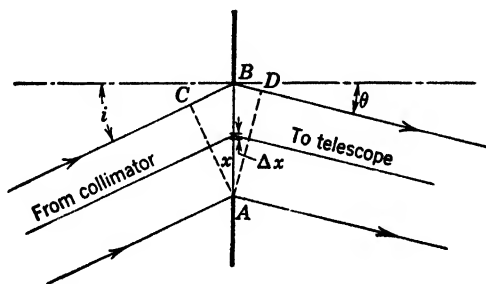


FIG. 11.1. Fraunhofer diffraction by a single slit.

fraction θ may be quite small. Since the optical paths from any point in the focal plane of a lens to the various points on a con-

jugate plane are equal, the paths to the left of AC , which is normal to the incident rays, are equal. Similarly, the paths to the right of AD , which is normal to the axis of the telescope, are also equal. Following Fresnel, let us consider the light at the focus of the telescope to be built up by the superposition of wavelets from the various elements of area of the slit, represented by Δx . The amplitude of each wavelet is taken proportional to Δx , and its phase relative to that of a wavelet from A is

$$\frac{2\pi x}{\lambda} (\sin i + \sin \theta)$$

This expression shows that the phase difference is proportional to the distance x , so that, if AB is divided into any number of strips of equal width Δx , the phase difference between the wavelet from any one and the next will be the same. Hence the vibration figure is a portion of a regular polygon, of which the closing side is the resultant amplitude. The angle between the first and last side will be the sum of all the phase increments or simply the total phase difference between the first and last wavelet

$$\Delta = \frac{2\pi e}{\lambda} (\sin i + \sin \theta)$$

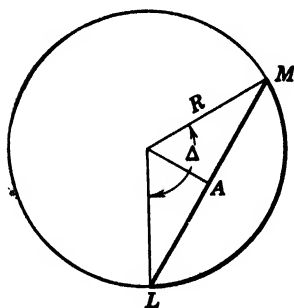


FIG. 11.2. Vibration curve for Fraunhofer diffraction by a single slit.

where e is the width of the slit. In the limit as Δx is made infinitesimal, the significant portion of the vibration polygon becomes a circular arc subtending the angle Δ , Fig. 11.2, since the angle between the radii to L and M is the same as the angle between the tangents at L and M which indicate the total phase difference. It is easy to see that the amplitude A of the diffracted wave is given by

$$A = 2R \sin \frac{\Delta}{2}$$

Now, if the telescope is rotated so that θ becomes equal to $-i$, the value of Δ approaches zero and the arc LM is transformed into a straight line of length LM , indicating the sum of wavelets which

are all in the same phase. Let A_0 represent this maximum amplitude at the principal focus of the telescope when it is in line with the collimator; then A_0 is equal to the length of the circular arc LM , or

$$A_0 = R\Delta$$

By eliminating R between these two equations, it is found that

$$A = A_0 \frac{\sin \frac{\Delta}{2}}{\frac{\Delta}{2}}$$

Since the intensity is proportional to the square of the amplitude, it follows that

$$I = I_0 \left[\frac{\sin \frac{\Delta}{2}}{\frac{\Delta}{2}} \right]^2$$

To examine the amplitude equation graphically, one may plot two functions, $y_2 = \sin \Delta/2$ and $y_1 = \Delta/2$, as in Fig. 11.3. The

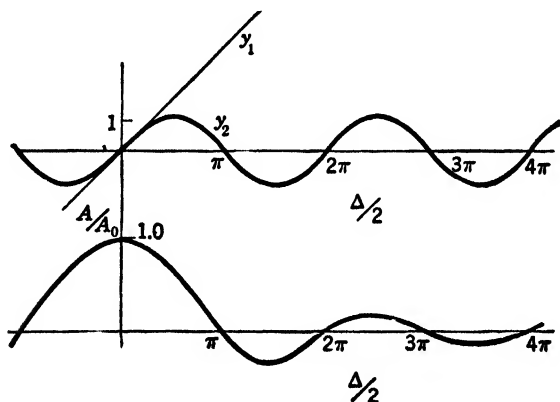


FIG. 11.3. Amplitude variation in the diffraction pattern by a single slit.

lower curve is a graph of their quotient $y_2/y_1 = A/A_0$. This quotient is zero whenever the numerator is zero, $\Delta/2 = \tau\pi$; $\tau =$

1, 2, 3, \dots , except when the denominator is zero as well, $\tau = 0$. In this case the quotient of the sine and the angle approaches unity. By squaring the ordinates one obtains a graph of the rel-

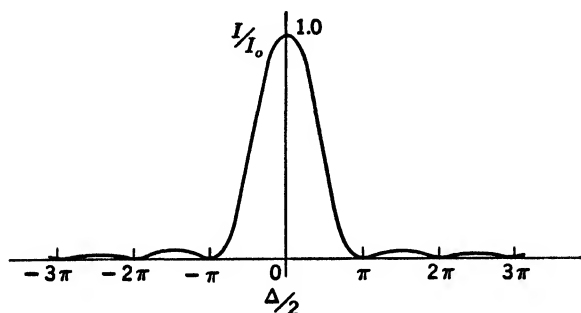


FIG. 11.4. Intensity distribution in the Fraunhofer diffraction pattern of a single slit.

ative intensity I/I_0 , which is drawn in Fig. 11.4. The intensity will be zero in the diffraction pattern when $\Delta/2 = \tau\pi$, or

$$\frac{\pi e}{\lambda} (\sin i + \sin \theta) = \tau\pi$$

or

$$\sin i + \sin \theta = \tau \left(\frac{\lambda}{e} \right)$$

One frequently makes $i = 0$ and finds that θ is so small that

$$\theta \doteq \tau \left(\frac{\lambda}{e} \right)$$

gives the directions in which minima are observed. If f is the focal length of the telescope or camera lens, the distances s from the center of the pattern to the various minima are

$$s = \tau \left(\frac{f\lambda}{e} \right); \quad \tau = 1, 2, 3, \dots$$

The minima are equally spaced at intervals $f\lambda/e$ except at the center, where the width of the central band is $2f\lambda/e$.

On the other hand, the maxima are not equally spaced, as can be shown by locating them by differentiation of the intensity equa-

tion. The resulting transcendental equation $\tan \Delta/2 = \Delta/2$ is usually solved graphically. As Δ increases, the roots approach $\Delta/2 = (\tau + 1/2)\pi$.

11.3 Circular Aperture

Since lenses and diaphragms usually have a circular boundary, diffraction by a circular aperture is of importance in the theory of resolving power. As before, the aperture may be divided into

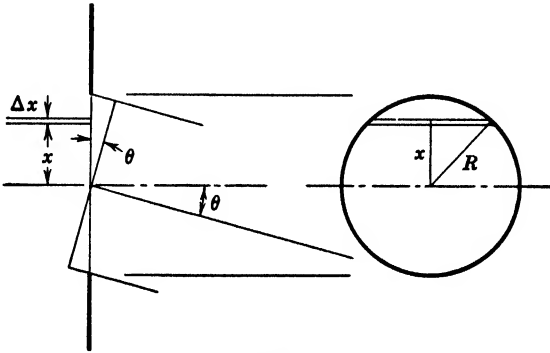


FIG. 11.5. Diffraction by a circular aperture showing the form of an element of area used in the integration procedure.

strips of equal width Δx , Fig. 11.5, for which the phase relative to the center of the aperture is

$$\frac{2\pi x}{\lambda} \sin \theta$$

assuming that $i = 0$. The phase changes by equal increments from one strip to the next. The amplitudes, however, are not equal, since the lengths of the strips are unequal, being given by $2\sqrt{R^2 - x^2}$. Thus the amplitudes of the wavelets da are proportional to $2\sqrt{R^2 - x^2} dx$. The value of A is given by the sum of the projections of these elementary amplitudes on A , Fig. 11.6. Thus

$$\begin{aligned} A &= 2 \int_{x=0}^{x=R} c \cdot 2\sqrt{R^2 - x^2} dx \cos \Delta \\ &= 4c \int_0^R \sqrt{R^2 - x^2} \cos \left(\frac{2\pi \sin \theta}{\lambda} x \right) dx \end{aligned}$$

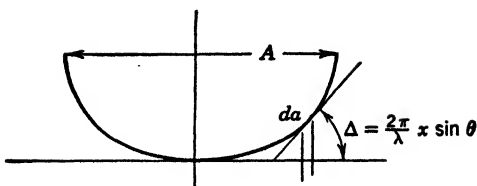


FIG. 11.6. Vibration curve for Fraunhofer diffraction by a circular aperture.

Let $u = x/R$ and let $m = (2\pi R \sin \theta)/\lambda$, then

$$A = 4cR^2 \int_0^1 \sqrt{1-u^2} \cos mu \, du$$

According to a mathematical theorem, the proof of which is omitted here:¹

$$\int_0^1 \sqrt{1-u^2} \cos mu \, du = \frac{\pi}{2} \frac{J_1(m)}{m}$$

where $J_1(m)$ is a Bessel function of the first order, the values of which may be found in mathematical tables such as *Tables of Functions* by Jahnke and Emde, page 157. The graph of $J_1(m)$

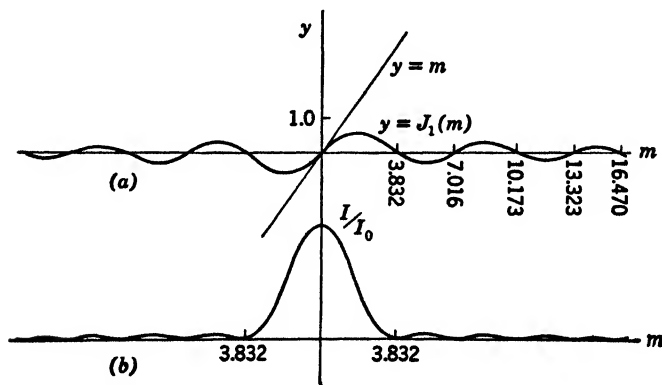


FIG. 11.7. (a) Graphs of the numerator and denominator of the factor $J_1(m)/m$. (b) Intensity distribution in the Fraunhofer diffraction pattern of a circular aperture.

resembles a damped sine wave, Fig. 11.7. Thus for a circular aperture

¹ Expand $\cos mu$ into a power series, integrate term by term with the aid of a trigonometric substitution, and compare the resulting series with that for $J_1(m)$.

$$A = 2\pi cR^2 \frac{J_1(m)}{m}$$

Let A_0 be the value of A when $\theta = 0$, or $m = 0$; then since

$$\lim_{m \rightarrow 0} \frac{J_1(m)}{m} = \frac{1}{2}$$

it follows that

$$A = 2A_0 \frac{J_1(m)}{m}$$

The intensity distribution is accordingly given by

$$I = I_0 \left(\frac{2J_1(m)}{m} \right)^2$$

The relative intensity I/I_0 is graphed against m in Fig. 11.7(b). The zero intensities lie at points corresponding to the roots of $J_1(m)$, except at $m = 0$. The first root, which determines the boundary of the central maximum, is at

$$m = 3.83 = \frac{2\pi R \sin \theta_1}{\lambda}$$

which gives the result

$$\sin \theta_1 = 1.22 \frac{\lambda}{2R}$$

Other roots of the Bessel function are indicated on the graph in Fig. 11.7(a). Because the Bessel function resembles a damped sine wave, the secondary maxima are considerably weaker than in the case of a slit.

11.4 Resolving Power

The resolving power of an instrument having a circular aperture of radius R is readily found from the last equation. Let θ be the angle subtended at the lens by two distant object points. The centers of the two diffraction patterns will also subtend the angle θ , Fig. 11.8. They will overlap more or less, according to the magnitude of the angle between object points as compared with the angular size of the central disk in the diffraction pattern. If the overlapping is too great, there will be a single maximum of in-

tensity instead of a double one, and the existence of two separate object points will not be perceived. According to Rayleigh, the images are just resolvable if the center of one is on the first minimum of the other. Consequently, the smallest angle resolvable is

$$\theta_m = 1.22 \frac{\lambda}{2R}$$

in radians.

If the object is not very far away, as in observations with a microscope, the result is somewhat different. If the index of re-

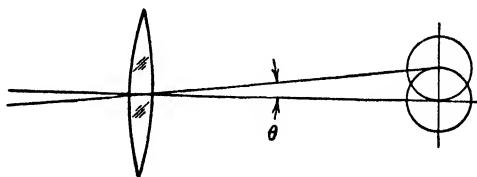


FIG. 11.8. Rayleigh's condition for the resolution of two adjacent diffraction disks.

fraction in the object space is n , the smallest linear separation between object points which may be resolved is

$$d_m = 1.22 \frac{\lambda}{2n \sin U}$$

where U is the angle subtended by the radius of the entrance pupil. The product $n \sin U$ is called the **numerical aperture**. This formula reduces to that just derived for θ_m when n is put equal to unity and the object is assumed to be far enough away so that $\sin U$ approximates $\tan U$. A derivation of the general formula may be found in Meyer's *Diffraction of Light, X-rays, and Material Particles*, pages 205–211.

11.5 The Double Slit

The phenomenon accompanying transmission of coherent waves by a double slit was partially discussed as an example of interference due to division of the wave front. However, the intensity variation can be completely accounted for only by applying diffraction theory. Even if a collimating and a projection lens are not used, the source and receiving screen are often so far from the

$$\Sigma x = A_1[1 + \cos \Delta + \cos 2\Delta + \cdots \cos (N - 1)\Delta]$$

$$\Sigma y = A_1[0 + \sin \Delta + \sin 2\Delta + \cdots \sin (N - 1)\Delta]$$

As before, the intensity distribution may be obtained by setting

$$I = (\Sigma x)^2 + (\Sigma y)^2 = (\Sigma x + i \Sigma y)(\Sigma x - i \Sigma y)$$

Each of these factors is a geometric series of N terms, the product of their sums being

$$I = A_1^2 \left(\frac{1 - e^{iN\Delta}}{1 - e^{i\Delta}} \right) \left(\frac{1 - e^{-iN\Delta}}{1 - e^{-i\Delta}} \right)$$

or

$$I = A_1^2 \left(\frac{1 - \cos N\Delta}{1 - \cos \Delta} \right) = A_1^2 \frac{\sin^2 N\Delta/2}{\sin^2 \Delta/2}$$

The factor A_1^2 has the form of the diffraction pattern of a single slit. The form of the second factor is best found by considering the graphs of the two functions $\sin N\Delta/2$ and $\sin \Delta/2$, which are shown in Fig. 11.14(a). Their quotient is zero whenever $\Delta/2 =$

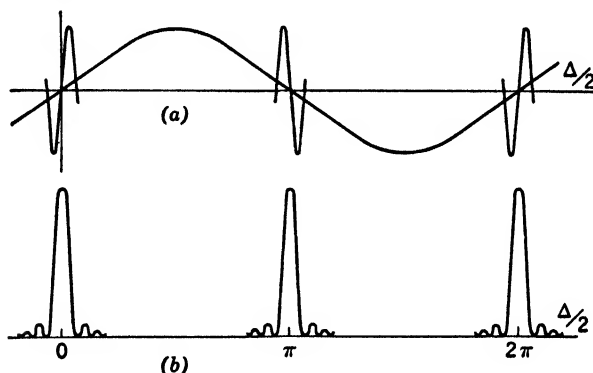


FIG. 11.14. (a) Graphs of the numerator and denominator of the principal factor in the amplitude equation for a grating. (b) The square of the quotient of the ordinates in (a).

$m\pi/N$, except when $\Delta/2 = m\pi$, m being any integer. In the second case, both numerator and denominator are zero, but the quotient can be evaluated by taking the ratio of the first derivatives of the numerator and denominator. This ratio is found to be N , so that $I = A_1^2 N^2$ when $\Delta/2 = m\pi$. These are the prin-

incipal maxima the location of which may be found by the condition

$$\frac{\Delta}{2} = \frac{\pi s}{\lambda} (\sin i + \sin \theta) = m\pi$$

or

$$s(\sin i + \sin \theta) = m\lambda$$

This formula is used in computing wave lengths from experimental values of i and θ as measured by a spectrometer. The value of m

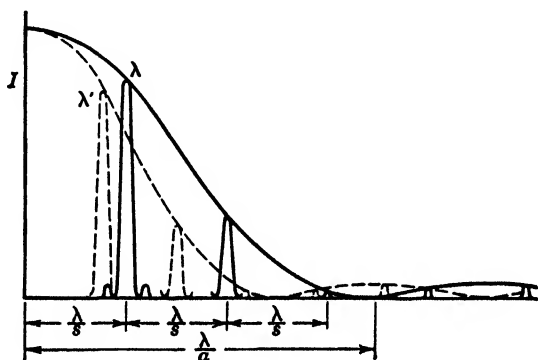


FIG. 11.15. Intensity distribution in the diffraction pattern of a grating consisting of N slits of width a and separation s . Graphs for two wave lengths are drawn to show the variation of intensity, dispersion, and resolving power with the order of diffraction.

denotes the “order” of the spectrum line. The same equation was found to hold for the maxima in the diffraction pattern of a double slit. In the case of a grating, however, the maxima are sharp (see Fig. 11.14), so that nearby wave lengths may be readily resolved if the grating has a sufficient number of lines. The secondary maxima between orders are relatively weak as compared with the principal maxima; in fact, they are imperceptible when the grating has many lines.

The intensities of the principal maxima of increasing order m vary in the manner indicated by A_1^2 , which is a function of i and θ . For grating elements which are narrow slits, this variation is like that observed in the diffraction pattern of a single slit. Thus the complete picture of the intensity distribution is approximately as shown in Fig. 11.15, where graphs of intensity for two wave lengths are indicated.

Modern gratings do not consist of a series of slits or of alternately reflecting and non-reflecting strips. Instead, smooth grooves are made by a diamond point on a surface of glass or of metal. These grooves must be regularly spaced to a very high degree of accuracy if the grating is to be satisfactory. Periodic errors in spacing as small as one millionth of an inch give rise to a disturbing pattern of "ghost" lines. Long-period variations in spacing give rise to groups of spurious lines in the vicinity of each genuine spectrum line. These are called **Rowland ghosts**. Those due to short-period fluctuations are often found in the extreme ultraviolet and are called **Lyman ghosts**. An explanation of the formation of ghost lines may be found in Meyer's *Diffraction of Light, Electrons, and Material Particles*, pages 174-185. The best modern gratings have been ruled by Wood and by Gale. The grating element in such a grating is represented by one period of its surface structure. Each elementary strip of infinitesimal width in each grating element is a source of a wavelet whose phase depends on the form of the groove. The wave obtained by summing up wavelets from all elementary strips in one element has the amplitude A_1 for this element. These wavelets are added as before to find the intensity curve for the grating. However, if the grooves are unsymmetrical in form, A_1^2 will not be a symmetrical curve and the relative intensities in the various orders will differ from those observed with a series of slits. By experimenting with various groove forms, R. W. Wood has succeeded in ruling gratings which diffract most of the radiant energy into a single order on one side of the central maximum. In this way, gratings have been made which give spectra whose intensities are not much inferior to those obtained with prisms.

11.7 Resolving Power of a Grating

As shown before, Fig. 11.14, the half width of a single principal maximum is represented by a change in $\Delta/2$ of π/N . The corresponding $d\theta$ is found from

$$d \frac{\Delta}{2} = \frac{\pi s}{\lambda} \cos \theta d\theta = \frac{\pi}{N}$$

so that

$$d\theta = \frac{\lambda}{Ns \cos \theta}$$

Note that $Ns \cos \theta$ is the complete width of the beam leaving the grating, which suggests an alternative derivation of $d\theta$ as the half width of the diffraction maximum by an aperture whose width is $Ns \cos \theta$. Applying Rayleigh's criterion for resolution, it is concluded that this value of $d\theta$ is the smallest angle between principal maxima of different λ that can just be resolved.

The dispersion D of a grating is obtained by differentiating the grating formula, thus:

$$m \, d\lambda = s \cos \theta \, d\theta$$

so that

$$D \equiv \frac{d\theta}{d\lambda} = \frac{m}{s \cos \theta}$$

Hence the increment in wave length that can just be resolved is

$$d\lambda_m = \frac{d\theta_m}{D}$$

or

$$d\lambda_m = \frac{\lambda}{Ns \cos \theta} \cdot \frac{s \cos \theta}{m} = \frac{\lambda}{Nm}$$

The resolving power is defined by the quotient of the wave length to the smallest change in wave length that can be detected. Therefore

$$R \equiv \frac{\lambda}{d\lambda_m} = Nm$$

A grating of 14,438 lines per inch and a ruled area of 6-in. width will have 86,628 lines, so that in the second order its resolving power will be 1.7×10^5 . At a wave length of 5100 Å one would be able to separate lines differing by only 0.03 Å with such a grating.

11.8 Fresnel Diffraction Phenomena

Fresnel diffraction phenomena are observed when spherical wave fronts are limited by obstacles or diaphragms. No lenses are used, as in observing the Fraunhofer diffraction phenomena previously considered. When the distances from source to obstacle and from obstacle to point of observation are increased indefinitely, the Fresnel phenomena merge into those of Fraunhofer as a limiting case. No unusual equipment is needed to observe Fresnel dif-

fraction, but the source must be a narrow line or point source, and the distances should be large enough (of the order of one meter) to obtain a fringe pattern whose structure is easily resolved by the unaided eye. Figure 11.16 is an example of a photographic record

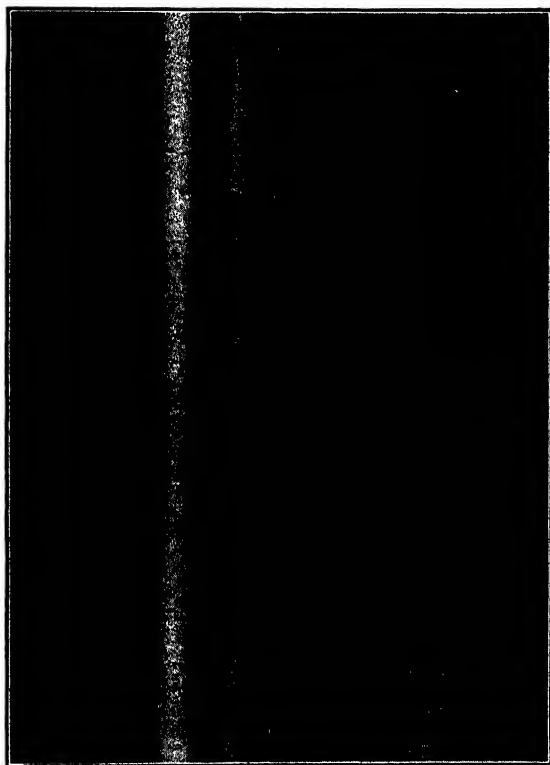


FIG. 11.16. Photograph of diffraction bands outside the shadow of a straight edge. Note the gradual decrease of intensity into the shadow of the obstacle.

of such a diffraction pattern. The first observations of diffraction were made under similar conditions by Grimaldi in 1650, and Newton describes several diffraction experiments in his *Opticks*, which was published in 1704.

Kirchhoff showed that the wave function at any point P , Fig. 11.17, may be expressed as an integral over any surface enclosing P and excluding the source Q . Each differential may be thought of as a Huygens wavelet originating at some element of area, as

dS . These are summed up, integrated, with due regard to their phase at P . As in Fresnel's theory, the amplitude of each wavelet is proportional to the element of area, so that $dA = k dS$, but Kirchhoff gives the following expression for Fresnel's arbitrary constant of proportionality:

$$k \equiv \frac{A}{2\lambda} \frac{\cos(n, r_1) - \cos(n, r)}{r_1 r}$$

In this expression A is the amplitude of the original wave at unit distance from Q , n is the normal to the surface S , the positive

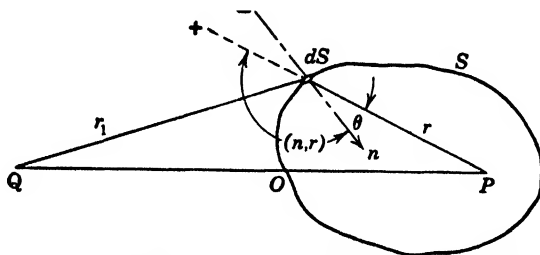


FIG. 11.17. The quantities appearing in Kirchhoff's solution of the wave equation.

direction being considered inward, r is the distance from P to the element of surface dS , and r_1 is the distance from Q to dS . There is a 90° phase difference between Fresnel's and Kirchhoff's solutions which will be ignored for the present, since it does not affect the intensity of the wave. In practice, the integral may be greatly simplified, since only the region near O contributes appreciably to the integral.

Consider a slit parallel to the y axis as shown in Fig. 11.18, with Q on one side and P on the other. The surface enclosing P will be taken as the xy plane, the rest of the surface being infinitely remote. Furthermore, only the integral over the slit need be considered, partly because of the infinite distance of the enclosing area on the right, and partly because of the absence of a back wave. Fresnel made the latter an arbitrary assumption, but it can be shown to follow directly from Kirchhoff's expression for Fresnel's k . Far to the right, the cosines of the angles (n, r) and (n, r_1) are equal, so that k is zero. The denominator is not a decisive factor. Although k varies inversely as the square of the distance, the area

over which the integral is taken increases as the square of the distance, so it is the cosine factor which reduces the back wave to zero. Moreover, experimental observation shows that the light intensity at P is not perceptibly changed unless the wave front

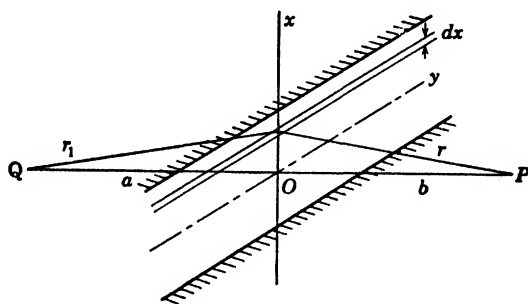


FIG. 11.18. Fresnel diffraction by a single slit.

near O is restricted, so that it should be allowable to reduce the area of integration to that of the slit.

Each elementary strip of width dx , Fig. 11.18, contributes a wavelet of amplitude $k dx$ and a phase relative to the strip through O of

$$\delta = \frac{2\pi}{\lambda} [(r_1 - a) + (r - b)]$$

By the use of the sagitta formula and by defining an auxiliary variable v to simplify the expression, one finds that

$$\delta = \frac{\pi}{\lambda} \left(\frac{1}{a} + \frac{1}{b} \right) x^2 = \frac{\pi}{2} v^2$$

The resulting wave at P may be obtained most readily from a vibration curve, Fig. 11.19, of which each element of length $k dx$ makes the proper phase angle with the wavelet from the pole O which is in the xy plane and on the line QP . The intensity at P is then proportional to the sum of the squares of total x and y components of all the elements $k dx$, so that

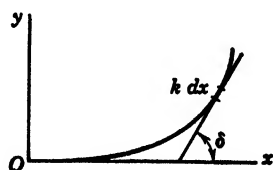


FIG. 11.19. A portion of the vibration curve for Fresnel diffraction.

$$I_P = \left(k' \int \cos \delta \, dx \right)^2 + \left(k' \int \sin \delta \, dx \right)^2$$

or

$$I_P = \left(k'' \int_{v_1}^{v_2} \cos \frac{\pi v^2}{2} \, dv \right)^2 + \left(k'' \int_{v_1}^{v_2} \sin \frac{\pi v^2}{2} \, dv \right)^2$$

where

$$v = \left(\sqrt{\frac{2(a+b)}{ab\lambda}} \right) x$$

The integrals which appear here are called **Fresnel integrals**. From tables of values of such integrals, one finds, in particular, that

$$\int_{-\infty}^{+\infty} \sin \frac{\pi v^2}{2} \, dv = \int_{-\infty}^{+\infty} \cos \frac{\pi v^2}{2} \, dv = 1$$

It follows that the factor $(k'')^2$ is equal to $A^2/2(a+b)^2$, because the intensity I_P must reduce to the value $A^2/(a+b)^2$ when there is no obstruction, or when $x_1 = -\infty$ and $x_2 = +\infty$. Therefore

$$I_P = \frac{A^2}{2(a+b)^2} \left\{ \left(\int_{v_1}^{v_2} \cos \frac{\pi v^2}{2} \, dv \right)^2 + \left(\int_{v_1}^{v_2} \sin \frac{\pi v^2}{2} \, dv \right)^2 \right\}$$

The quantity in brackets can best be evaluated with the aid of a Cornu spiral, which is a graph of

$$\eta = \int_0^v \sin \frac{\pi v^2}{2} \, dv \quad \text{against} \quad \xi = \int_0^v \cos \frac{\pi v^2}{2} \, dv$$

The properties of the Cornu spiral, Fig. 11.20, may be found as follows:

Since

$$d\xi = \cos \frac{\pi v^2}{2} \, dv \quad \text{and} \quad d\eta = \sin \frac{\pi v^2}{2} \, dv$$

an element of length along the spiral is

$$ds = \sqrt{d\xi^2 + d\eta^2} = dv$$

so that, with the proper choice of origin,

$$s = v$$

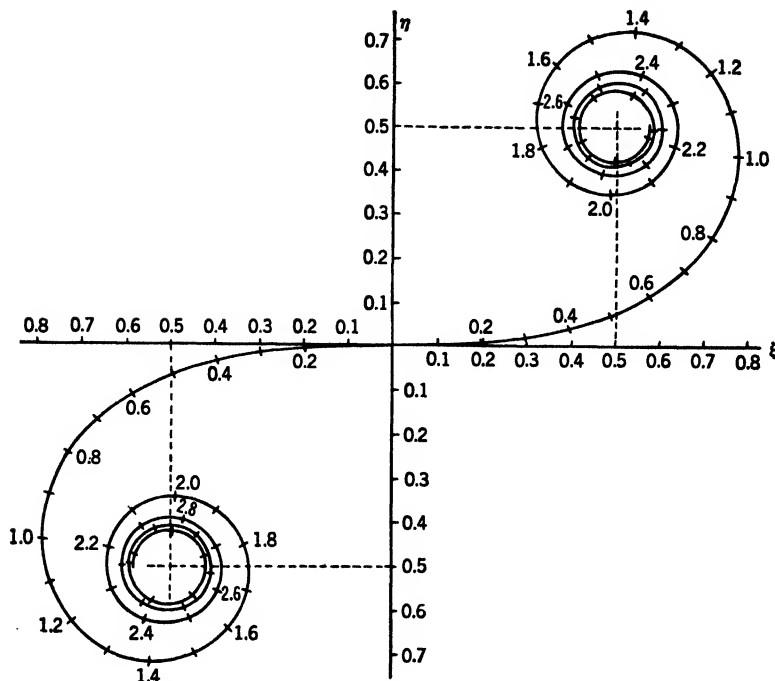


FIG. 11.20. Cornu spiral. A graph of the Fresnel integrals.

Thus the value of v is the distance s along the spiral measured from the origin. Moreover, the slope at any point

$$\tan \tau = \frac{d\eta}{d\xi} = \tan \frac{\pi v^2}{2}$$

so that

$$\tau = \frac{\pi v^2}{2} = \delta_m$$

where δ_m is the phase difference between the wavelet from a strip in the slit at a distance x from the pole of the wave and the wavelet from O . Finally, the radius of curvature

$$\rho = \frac{ds}{d\tau} = \frac{1}{\pi v} = \frac{1}{\pi s}$$

decreases as s increases, yielding a non-intersecting double spiral as shown in Fig. 11.20. The exact numerical values of the Fresnel

integrals may be computed by means of a series expansion and have been tabulated in various places.² The spiral is found to converge toward the points

$$\xi = +\frac{1}{2}, \quad \eta = +\frac{1}{2}; \quad \text{and} \quad \xi = -\frac{1}{2}, \quad \eta = -\frac{1}{2}$$

The intensity variation near the shadow of a straight edge, Fig. 11.21, may be found from the appropriate chords of the Cornu spiral. For any point P , the upper limit of integration is infinity, corresponding to A in Fig. 11.22, but the lower limit x_1 , and there-

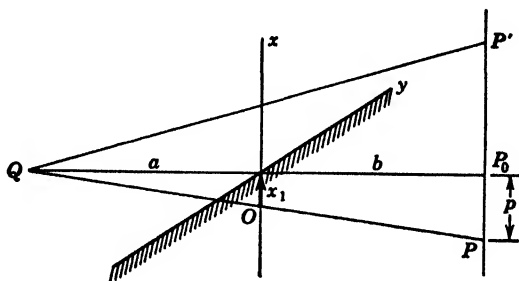


FIG. 11.21. Fresnel diffraction at a straight edge.

fore v_1 , varies with the location of P . Let P_0 be a point located at the geometric boundary of the shadow. Then, for a point P at a distance p below P_0 , the value of

$$x_1 = \frac{a}{a+b} p$$

is found by similar triangles, Fig. 11.21. The distance s_1 along the Cornu spiral to the lower limit of the chord is then

$$s_1 = v_1 = \sqrt{\frac{2(a+b)}{\lambda ab}} \left(\frac{ap}{a+b} \right) = \left(\sqrt{\frac{2a}{\lambda b(a+b)}} \right) p$$

This is positive for positive p , or when P is inside the shadow, and is negative if P is outside the shadow, as P' , Fig. 11.21.

Imagine a rubber band with one end fastened at the point A ($\frac{1}{2}, \frac{1}{2}$), while the other end moves along the spiral. The length

² MONK, *Light Principles and Experiments*, McGraw-Hill, 1937, page 440; JENKINS and WHITE, *Fundamentals of Physical Optics*, McGraw-Hill, 1937, page 189.

of this band represents the amplitude of the wave at P . Inside the shadow, i.e., positive v_1 , the intensity diminishes smoothly, whereas outside the shadow there are maxima and minima, as one can see by a study of Fig. 11.22. These are approximately on the 45° line AA' . Since the spiral cuts this line nearly at right angles, one may write the conditions for maximum and minimum intensity

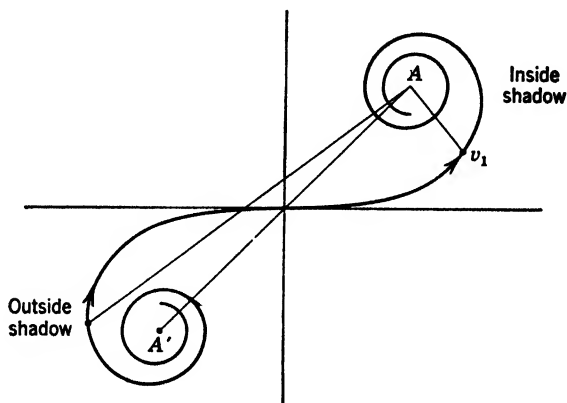


FIG. 11.22. Application of Cornu's spiral to Fresnel diffraction at a straight edge.

in terms of the slope angle τ of the spiral along this 45° line. Thus for the maxima

$$\tau = 2\pi m + \frac{3\pi}{4}$$

From this one finds the locations of the corresponding points P by writing

$$2\pi m + \frac{3\pi}{4} = \frac{\pi v^2}{2} = \frac{\pi(a+b)}{\lambda ab} x^2 = \frac{\pi a}{\lambda b(a+b)} p^2$$

so that

$$p_{\max} = \sqrt{\frac{\lambda b(a+b)}{2a} \left(4m + \frac{3}{2}\right)}; \quad m = 0, 1, 2, 3 \dots$$

Similarly, one finds that the minima occur when

$$\tau = 2\pi m + \frac{7\pi}{4}$$

or at the distances

$$p_{\min} = \sqrt{\frac{\lambda b(a+b)}{2a} \left(4m + \frac{7}{2}\right)}$$

The corresponding intensities are proportional to the squares of the chords of least and greatest length, measured from A to the convolutions in the lower left-hand quadrant. They are con-

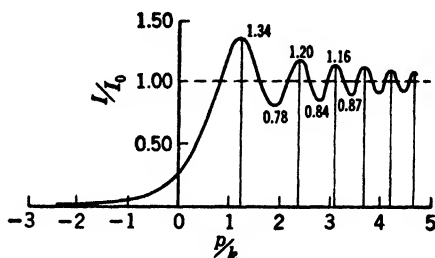


FIG. 11.23. Intensity distribution in the Fresnel diffraction pattern at the shadow of a straight edge.

veniently expressed in terms of the intensity of the unobstructed wave which is proportional to the square of AA' .

If we let the factor $k \equiv \sqrt{\frac{\lambda b(a+b)}{2a}}$, the following values are found for the locations and magnitudes of the maxima and minima.

Maxima: $I_1 = 1.34$ $I_2 = 1.20$ $I_3 = 1.16$... $I = 1$

at $p_1 = 1.225k$ $p_2 = 2.345k$ $p_3 = 3.082k$...

Minima: $I'_1 = 0.78$ $I'_2 = 0.84$ $I'_3 = 0.87$... $I' = 1$

at $p'_1 = 1.871k$ $p'_2 = 2.739k$ $p'_3 = 3.391k$...

These are indicated on the intensity graph in Fig. 11.23, a photograph of the pattern being reproduced in Fig. 11.16.

The intensity distribution in the Fresnel diffraction pattern of a single slit is obtained by taking squares of the chords of a constant length of the Cornu spiral. The ends of the chord are located at

$$v_1 = \sqrt{\frac{2(a+b)}{\lambda ab}} x_1 \quad \text{and} \quad v_2 = \sqrt{\frac{2(a+b)}{\lambda ab}} x_2$$

where x_1 and x_2 are related to the location of P and the width of the slit s as shown in Fig. 11.24. Thus

$$x_1 = \frac{a}{a+b} p - \frac{s}{2} \quad \text{and} \quad x_2 = \frac{a}{a+b} p + \frac{s}{2}$$

Three cases may be distinguished. (1) If the width of the slit is so small that $v_2 - v_1$ is of the order of magnitude of 0.1 or less, there is merely a diffusion of light behind the slit, with no ob-

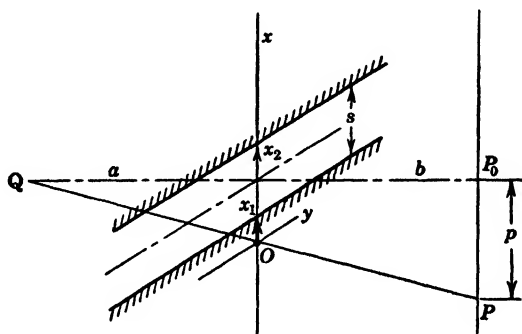


FIG. 11.24. Fresnel diffraction at a single slit, showing the limits of integration x_1 and x_2 for a point P .

servable maxima and minima. (2) If the width is such that $v_2 - v_1$ is about 1.0, there will be maxima and minima only in the shadow region where v_2 and v_1 have the same sign. The pattern resembles qualitatively the Fraunhofer diffraction pattern, but the intensity minima are not zero. (3) If the width is such that v_2 and v_1 have opposite signs at the maxima and minima, the diffraction bands lie within the geometrical boundary of the central illuminated region. These are called "inner bands," and as the slit is widened they approach in appearance the fringes lining the shadow of two straight edges.

Note that, on the contrary, the Fraunhofer diffraction pattern of a slit becomes merely narrower as the slit is widened, without any change in the nature of the diffraction pattern. In other words, one may use the same equation or graph, with merely a change in scale, to describe the intensity variation in the Fraunhofer case.

While we have been assuming a point source Q , one usually takes an illuminated slit or narrow luminous filament as the source of light. If this is parallel to the straight edge or slit, the diffraction

fringes for each point of the source will fall on the same straight line. Thus the fringes will be intensified without any loss in distinctness.

A narrow obstacle may be treated as the complementary of a slit. The two chords of the Cornu spiral $A'v_1$ and $A'v_2$ respectively give the amplitudes and phases of the waves coming around the two sides of the obstacle. They must be added vectorially to obtain the resultant amplitude, whose square is the intensity. It is readily seen that there will be equally spaced bands inside the shadow where v_1 and v_2 are of opposite sign. Outside the shadow, where the values of v have the same sign, the intensity follows approximately the same fluctuations as with a straight edge.

The paradoxical situation of a maximum intensity in the center of the shadow of an opaque obstacle is seen to follow from the Cornu spiral regardless of the width of the obstacle. One finds that, in the center of the shadow, v_1 and v_2 are equally distant on both sides of the origin. Consequently, the chords $A'v_1$ and v_2A are always parallel, giving rise to a maximum sum or to "constructive interference." The magnitude of the sum, however, diminishes as the obstacle is made wider.

Similarly, in the case of a circular obstacle, there is a bright spot in the center of its shadow if a point source is used. The spot is surrounded by circular bands of smaller intensity. Outside the shadow, circular diffraction bands are seen.

The Fresnel diffraction pattern of a circular aperture has the same general characteristics as that of a slit except that the diffraction fringes are circular and the scale of the diffraction pattern is somewhat larger. As the diameter of the aperture increases, one finds a similar variation from diffusion through "outer bands" to "inner bands." The intensity at the center of the pattern is alternately a maximum and a minimum. This is readily seen to follow from a consideration of the phases of wavelets from the various zones of the incident wave front. Let b_0 , Fig. 11.25, represent the distance from the point of observation P to the nearest point or "pole" of the spherical wave. Following Fresnel's procedure, one may subdivide the wave front into circular zones about the pole O , as follows;

First zone:	$b_0 <$	$b < b_0 + \lambda/2$
Second zone:	$b_0 + \lambda/2 <$	$b < b_0 + 2\lambda/2$
Third zone:	$b_0 + 2\lambda/2 <$	$b < b_0 + 3\lambda/2$
n th zone:	$b_0 + (n - 1)\lambda/2 <$	$b < b_0 + n\lambda/2$

The contributions to the amplitude at P from each of these zones decrease in value because of the cosine factor in Kirchhoff's k and alternate in phase because of the increase in distance by intervals of half a wave length. Thus an aperture transmitting an even number of zones will give a minimum intensity at P , whereas an aperture transmitting an odd number of zones will give a maximum. These maxima and minima will be most pronounced if the aperture transmits a small number of zones, for then the amplitudes of the wavelets are larger. In fact, the first zone alone gives nearly twice

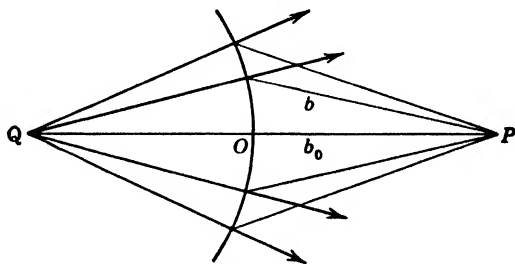


FIG. 11.25. Division of spherical wave front into Fresnel zones. The distance b to the boundary of each zone increases by steps of half a wave length.

the amplitude, or four times the intensity, due to the entire wave, while two zones give an amplitude of nearly zero.

It is evident that the number of zones transmitted by a given aperture depends on the distance b_0 as well as on the wave length. To find the radius r_n of the aperture transmitting n zones, one finds, by the theorem of Pythagoras, that

$$r_n^2 = b^2 - b_0^2 = n\lambda b_0 + n^2 \frac{\lambda^2}{4} = n\lambda b_0$$

A *zone plate* is a device which transmits alternate zones. Since the wavelets from these zones are in phase, it gives a high concentration of intensity at the "focus" P , and it may be used as a lens to concentrate light or to create images.

11.9 Kirchhoff's Formulation of Fresnel's Theory

It has been indicated that Kirchhoff derived an expression for the solution u_P of the general wave equation at a point P in terms of a surface integral over a surface enclosing P but not also including the source Q , Fig. 11.17. Let n represent the normal to

such a surface, considered positive inward, r the distance from the point P to an element of area dS on the surface. Let u , a function of time and space coordinates, represent the wave travelling from the source Q outside the surface S . Then Kirchhoff's formula³ reads

$$4\pi u_P = \iint \left\{ \frac{\partial}{\partial r} \left[\frac{u \left(t - \frac{r}{V} \right)}{r} \right] \cos(n, r) - \frac{1}{r} \frac{\partial}{\partial n} \left[u \left(t - \frac{r}{V} \right) \right] \right\} dS$$

where $u(t - r/V)$ is the expression for the wave with its time variable replaced by $t - r/V$ to take into consideration the phase retardation due to the time of propagation r/V from S to P . The equation expresses a generalized Huygens principle, which states that the effect u_P at any point P may be considered to equal the superposition of effects propagated with the velocity V from the elements dS of any surface enclosing P .

The mathematical operations are simpler if one writes the expression for the wave u in the complex form. For the special case of an incident plane wave this becomes

$$u = u_0 e^{i\omega t - ikr_1}$$

where $\omega = 2\pi/T$ and $k = 2\pi/\lambda$, while r_1 is the distance from any convenient coordinate plane of constant phase, a wave front, since Q is now at infinity. One may consider the real part of this complex expression as the physically significant part, or he may advantageously perform all mathematical operations, using complex quantities, and finally determine the intensity of the wave by multiplying the resulting u_P by its complex conjugate. Since $e^{i\omega t}$ is a factor which does not enter into the operations of differentiation and integration, it may just as well be omitted. The remaining complex expression for u or u_P gives the amplitude and phase of the wave.

If one has an opening in an opaque plane screen, he may take the enclosing surface S just back of this plane and then closing around P at infinite distances. Only the integration over the opening contributes significantly to u_P . One finds that

$$u_P = \frac{u_0}{4\pi} \iint \left[\frac{\partial}{\partial r} \left(\frac{e^{-ik(r+r_1)}}{r} \right) \cos(n, r) - \frac{1}{r} \frac{\partial}{\partial n} e^{-ik(r+r_1)} \right] dS$$

³ For example, see DRUDE, *Theory of Optics* (transl. by Millikan and Mann), Longmans, Green, 1920, page 179.

With perpendicularly incident plane waves, r_1 may be set equal to zero except when finding the rate of change along the normal, the reference plane being taken in the plane of the aperture. Carrying out the differentiations and letting $\theta = \pi - (n, r)$ represent the angle between the normal to the surface S and the line from dS to P , one may reduce this to the form

$$u_P = \frac{iu_0}{4\pi} \iint \left[\frac{2\pi}{\lambda} \frac{e^{-ikr}}{r} (1 + \cos \theta) - i \frac{e^{-ikr}}{r^2} \cos \theta \right] dS$$

since $\partial/\partial n = \partial/\partial r_1$ in this case, and $\cos (n, r) = -\cos \theta$. In the case of short optical waves with a wave length much less than r , one may neglect the second term of the integrand. Applications of this formula would repeat many of the results already developed by Fresnel's theory, so they will be omitted here.

The Kirchhoff theory is superior to Fresnel's in that it contains no arbitrary constants of proportionality, making it possible to compute absolute intensities of diffraction, not just the distributions of intensity over the pattern. Furthermore, the phase of the diffracted wave is correct instead of being in error by $\pi/2$. However, if one is satisfied with just the intensity distributions, the two theories give essentially the same results except at large angles of diffraction and at small distances, and the Fresnel theory is simpler to apply.

11.10 Diffraction of Microwaves

The development of sources and detectors of electromagnetic waves having a wave length of a few centimeters has made it possible to study diffraction at distances which are not great compared with the wave length. Measurements have even been made in the plane of the aperture.⁴ These experimental results show that the amplitude is not constant over the aperture, contrary to the usual assumption in developing the theory of diffraction. In fact, the fluctuations in amplitude are greatest in the plane of the aperture. The polarization and phase of the waves are also found to vary over the aperture. These facts point to the need for a revision of the entire theory of diffraction, which has not yet been carried out. The new theory should, however, asymptotically approach, or at least approximate, the Kirchhoff-Fresnel theory as one approaches

⁴ ANDREWS, *Physical Review*, **71**, 777 (1947).

the usual optical conditions for observing diffraction. As long as the apertures have dimensions of many wave lengths and the distances of observation are an even greater number of wave lengths, the Fresnel theory accounts for the principal facts about diffraction of radiations ranging from x-rays of a few angstrom units in wave length (**Kellstrom**) up to microwaves with wave lengths of several centimeters (**Andrews**). In fact, Andrews shows that, even when the conditions of the experiment deviate as greatly from those of optics as in his investigations, the results computed by the Kirchhoff formula are still sufficiently precise to be useful. One may consider this to be remarkable in view of the faulty assumptions on which it is based. However, it must be recognized that, when the wave length is not small compared with the size of the aperture and compared with the distance at which the diffraction pattern is studied, the Kirchhoff and Fresnel theories are not entirely dependable and should be used only as convenient approximations.

11.11 Diffraction of X-rays by Crystals

Crystals are characterized by a periodic structure in three-dimensional space. Let a , b , and c be the respective periods along the three crystallographic axes. Although these axes are not always at right angles, only the special cases when this is true will be considered here. A **space lattice** is made up of identical points or entities whose Cartesian coordinates are integral multiples of a , b , and c respectively.

When x-radiation passes through such a lattice, coherent wavelets are diffracted from each array of identical points. It is our problem to find the relations which must exist between the direction of propagation and the wave length in order that the wavelets reinforce each other constructively. Laue solved this three-dimensional diffraction problem by considering the simultaneous solution of three linear diffraction problems. Bragg, however, showed that the same final result may be obtained by two steps, as will be shown in the following. He recognized the fact that all the identical points in the three-dimensional space lattice may be considered to lie on equally spaced, parallel planes. There are many possible families of such planes, but only those with a relatively high density of points are significant. The situation is analogous to the many ways in which one may draw equally spaced,

parallel lines through an array of points having regular periodic spacing in two dimensions. Consider, for example, the various rows of plants that one may see on a regularly planted cornfield. Each family of parallel rows is characterized by a slope and a perpendicular interval between members. The same is true of crystal planes.

Crystal planes may be defined by their intercepts on the three crystallographic axes, or in terms of the reciprocals of these intercepts. For example, a plane which passes through m_1a on the x axis, m_2b on the y axis, and m_3c on the z axis may be designated by the integers $m_1m_2m_3$. These integers may also be used to specify the entire family of parallel planes. In this case, any common factor is divided out, leaving the three smallest integers which define the ratio of the intercepts. It is more common, however, to designate a family of planes by three integers hkl which are respectively proportional to the reciprocals of m_1 , m_2 , and m_3 . Thus

$$h:k:l = \frac{1}{m_1} : \frac{1}{m_2} : \frac{1}{m_3}$$

The reciprocals on the right are multiplied by a common factor to obtain the three smallest integers having no common factor other than unity. Such values of h , k , and l are called the **Miller indices** of a family of planes. For example, 100 planes are parallel to the y and z axes; 110 planes are parallel to the z axis and cut the x and y axes at equal multiples of a and b ; etc.

To show that the members of any family of planes are equally spaced, and to derive a formula for this spacing, one writes first the equation of one of the planes in the intercept form

$$\frac{x}{m_1a} + \frac{y}{m_2b} + \frac{z}{m_3c} = 1$$

Converting this into the direction cosine form, one finds the perpendicular distance P from the origin to the plane by the equation

$$P = Lx + My + Nz$$

where L , M , and N are the direction cosines of the normal to the plane. A comparison of the two equations shows that

$$L:M:N = \frac{1}{m_1a} : \frac{1}{m_2b} : \frac{1}{m_3c} = \frac{h}{a} : \frac{k}{b} : \frac{l}{c}$$

Since $L^2 + M^2 + N^2 = 1$, it follows that the distance P to a plane of the proper slope passing through any point x, y, z is given by

$$P = \frac{h\left(\frac{x}{a}\right) + k\left(\frac{y}{b}\right) + l\left(\frac{z}{c}\right)}{\sqrt{\frac{h^2}{a^2} + \frac{k^2}{b^2} + \frac{l^2}{c^2}}}$$

Since $x/a, y/b$, and z/c are the coordinates of the point in terms of the unit distances a, b , and c , it is convenient to write them respectively as x_1, y_1 , and z_1 . For example, a given set of Miller indices such as 111 defines planes with

$$P = \frac{x_1 + y_1 + z_1}{\sqrt{\frac{1}{a^2} + \frac{1}{b^2} + \frac{1}{c^2}}}$$

Actual lattice planes must pass through points x_1, y_1, z_1 whose coordinates are any three integers. It is quite evident that various values of P will be found, but that these will be discrete, changing in value by steps of

$$\Delta P = \frac{1}{\sqrt{\frac{1}{a^2} + \frac{1}{b^2} + \frac{1}{c^2}}}$$

This is accordingly the distance d between successive planes in the family 111. Similarly, for any hkl plane, one finds that

$$d = \frac{b}{\sqrt{\frac{h^2}{A^2} + k^2 + \frac{l^2}{C^2}}}$$

where $A = a/b$ and $C = c/b$ are called the **axial ratios** of the crystal system. In the cubic system $A = C = 1$, and one may write:

$$d = \frac{a}{\sqrt{h^2 + k^2 + l^2}}$$

Formulas for the interplanar distances in crystals whose axes are not at right angles to each other may be found in Clark's *Applied X-rays*, 3d ed., page 283.

The diffraction problem is easily solved in two steps by employing the concept of scattering centers in families of equidistant planes. First, there is constructive interference between wavelets from all points in any one plane, regardless of their separation and the wave length of the radiation, if the law of reflection is satisfied. In x-ray work it is customary to measure angles of rays with the plane itself instead of with its normal, as shown in Fig. 11.26. Such angles are called **glancing angles**, and our first requirement for constructive interference is that $\theta' = \theta$. The second require-

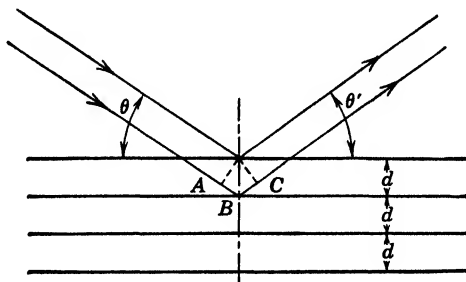


FIG. 11.26. Bragg's solution of the problem of diffraction by a crystal.

ment is that all such "reflections" from the various planes of a given family must be in the same phase. This will be true if the path difference $AB + BC$, Fig. 11.26, is a whole number of wave lengths. Since $AB = BC = d \sin \theta$, it follows that the intensity maxima occur when

$$2d \sin \theta = n\lambda$$

where n , an integer, is the order of diffraction. This equation is called Bragg's formula. The maxima are very sharp because of the large number of scattering elements in a crystal of ordinary size. Because $\theta' = \theta$, the x-rays are often said to be reflected from the planes of atoms, but this is not actually the case. The fact that each wave length requires a different angle

$$\theta = \sin^{-1} n \left(\frac{\lambda}{2d} \right)$$

for its "reflection" shows that the process is not true reflection at all.

Bragg's formula, with appropriate expressions for d , is widely used in the analysis of crystal structure by means of x-ray dif-

fraction. One may determine the periodicities a , b , and c from the directions of three suitably chosen diffraction maxima. However, a complete analysis also includes a determination of the distribution of atoms in a unit cell of the crystal. This does not affect the directions of the maxima, but does affect their relative intensities, sometimes reducing certain ones to zero. An analogous situation exists in the pattern of an ordinary optical grating, in which the periodicity of the rulings determines the directions of the maxima, but the groove form determines the distribution of intensity in the various orders. The use of intensity measurements in crystal structure analysis is important, but will not be discussed further here.

Bragg's formula is also useful in measuring wave lengths by the methods of x-ray spectroscopy. Accurate measurements require that the index of refraction of the crystal be taken into consideration. The index μ is slightly less than unity in the x-ray region and is often written as $\mu = 1 - \delta$. One may designate the glancing angle on all planes inside the crystal by θ'' . Then by Snell's law

$$\frac{\cos \theta}{\cos \theta''} = \mu$$

The wave length inside the crystal λ'' is related to that outside by

$$\frac{\lambda}{\lambda''} = \mu$$

Thus Bragg's law for the crystal interior, where most of the diffraction takes place, is

$$n\lambda'' = 2d \sin \theta'' = 2d\sqrt{1 - \cos^2 \theta''}$$

Substituting the expressions for λ and θ outside the crystal, where the measurements are made, one finds that

$$n \frac{\lambda}{\mu} = 2d\sqrt{1 - \frac{\cos^2 \theta}{\mu^2}}$$

or

$$n\lambda = 2d\sqrt{\mu^2 - \cos^2 \theta}$$

or

$$n\lambda = 2d\sqrt{\mu^2 - 1 + \sin^2 \theta}$$

Since μ is unity within a few millionths, $\mu^2 - 1 \doteq -2\delta$ so that

$$n\lambda = 2 \left(d \sqrt{1 - \frac{2\delta}{\sin^2 \theta}} \right) \sin \theta$$

Because of the smallness of the second term under the radical, one may use the first two terms of the binomial expansion of the square root, finding that

$$n\lambda = 2d \left(1 - \frac{\delta}{\sin^2 \theta} \right) \sin \theta$$

or

$$n\lambda = 2d \left(1 - \frac{\delta}{\lambda^2} \frac{4d^2}{n^2} \right) \sin \theta$$

by the use of the approximate value of $\sin \theta$ in the small correction term in the parentheses. Bragg's formula may finally be written in the form

$$n\lambda = 2d_n \sin \theta$$

where d_n is an apparent spacing which includes a correction for refraction. Thus a slightly different value of d_n is used for each order of diffraction, as noted in Experiment 22, in which the standard values for calcite are listed.

PROBLEMS

1. Compute the locations and relative intensities of the first maximum and the first minimum in the diffraction pattern of a straight edge on a screen 2 meters from the edge with the light source very far away on the other side, the wave length being (a) 5500 Å, and (b) 1 cm.

2. Prove the formula for the locations of the dark bands in the Fraunhofer diffraction pattern of a single slit, and explain why the same formula gives the locations of the bright bands in the case of Young's experiment when the single slit is converted into two slits by a broad central opaque partition, leaving two equally narrow slits.

3. Describe the Fraunhofer and the Fresnel diffraction patterns for a slit whose width is varied from zero to a very large value. Distinguish between the several cases, and give a simplified explanation of the diffraction patterns with the aid of appropriate vibration curves.

4. At least how large must a lens be in order to focus light of 5000 Å to a spot 1 μ in diameter if the focal length of the lens is 200 mm?

5. At least how far apart must two small lights be placed in order that a person 5 miles away may distinguish them as separate objects when looking

through a telescope whose objective has a diameter of 2 in.? (Use wave length = $1/50,000$ in.)

6. What is the size of the image of a star produced by the objective of a telescope whose diameter is 36 in. and whose focal length is 30 ft, assuming a perfect lens and a wave length of $1/50,000$ in.?

7. The objective of a battery commander's telescope has a diameter of 1.78 in. (a) How small is the structural detail which may be resolved with this telescope at a distance of 2 miles? (b) At least what magnification is required?

8. How large a mirror is needed to focus microwaves having a wave length of 2 cm onto a circular area 10 cm in diameter at a distance of 10 meters?

9. What limit does diffraction set on the resolving power of the human eye if the diameter of the pupil is 2.2 mm and the wave length of light is 5500 \AA ? Express the result in terms of the smallest resolvable visual angle in minutes.

10. If the resolving power of a photographic film is 50 lines per millimeter, at what relative aperture will it begin to be possible to observe a decrease in resolution due to diffraction as the lens is stopped down? Let wave length equal 5000 \AA .

11. Derive the intensity distribution for a grating of N slits of width a spaced at intervals b and interleaved by N additional slits of width $a/2$, which divide the original intervals in the ratio of 1:3.

12. The objective of a certain telescopic sight has a diameter of 3 in. (a) How small is the structural detail which may be resolved at a distance of 1 mile? (b) If the unaided eye can resolve points subtending 1 minute or 2.9×10^{-4} radian, what minimum magnification is required to utilize the maximum resolving power?

13. A grating having 14,440 lines per inch is placed 1 meter from a scale on the perpendicular to the scale at its midpoint. If mercury light of wave length 5461 \AA is passed through a slit at the center of the scale, where will the first order diffraction image be seen on looking through the grating at the scale?

14. The separation of the Zeeman components of the sodium lines is about 0.025 \AA when the field strength is 15,000 gauss. How would you determine the suitability of a grating spectrograph for directly observing this effect?

15. (a) In choosing a prism spectrograph to resolve the deuterium and hydrogen alpha lines (separation 1.8 \AA at 6562 \AA), what size prism would be needed if $dn/d\lambda = 1500$ and one sacrifices half the maximum resolving power for a suitable intensity? (b) What focal length lens would give a 0.5-mm separation of these lines? Assume a 60° prism at minimum deviation and with an index of refraction of 1.600.

16. Compute the glancing angles at which one finds diffraction maxima of the first and second order when x-rays of 1.54 \AA wave length are reflected from a cleavage plane of calcite for which $d = 3.029 \text{ \AA}$.

17. What are the angles of diffraction of x-rays having a wave length of 1.92 \AA diffracted by the 100, 110, and 111 planes of rock salt? Lattice constant = 5.628 \AA .

Chapter 12. Polarization and Double Refraction

12.1 Polarized Light

When light is transmitted by a tourmaline crystal plate or by a sheet of Polaroid, its nature is changed, as can be shown by inserting and rotating a second plate about an axis that coincides with the direction of the light beam. If one measures angles of rotation ϕ from the position in which the transmitted intensity is greatest, it is found that the intensity I varies as the square of the cosine of ϕ

$$I = I_m \cos^2 \phi$$

Light behaving in this way is said to be **plane polarized**. The intensity formula is known as **Malus' law**, as it was discovered empirically in 1810 by Captain Étienne Malus. When ϕ is 90° or 270° , the intensity drops to zero if the polarization is complete; otherwise one observes only a minimum. If the light is unpolarized, no change in intensity is observed when it is tested with a single plate in this way. This and similar experiments are readily explained only if light waves are transverse. At one time, when light waves were thought to consist of elastic waves in the ether, the transverse nature of the waves was considered to be practically impossible, since it required the all-pervading ether to have the elastic properties of a solid. However, it is now well known that light is a member of the family of electromagnetic waves, and that these are transverse follows directly from Maxwell's theory, as will be shown presently.

Light waves coming directly from a source usually vibrate impartially in all directions at right angles to the wave normal. When light is plane polarized, the components of the vibrations in some plane are transmitted or reflected to the exclusion of the other components. The eye is unable to detect such a selection of components. Although the intensity is reduced to one half in the most efficient polarizer, its rotation causes no change in intensity. However, if a second polarizing device operates on the polarized light, it will transmit or reflect only the cosine components in its plane of vibration. Since the intensity varies as the square of the amplitude, Malus' law follows at once.

12.2 Electromagnetic Theory

Of the many methods by which light may be polarized, the most common is oblique specular reflection from the surface of an electrically insulating medium. The conditions for this effect and the degree of polarization obtained are readily derived by the use of the electromagnetic theory of light. Let E , whose components are X, Y, Z , represent the electric field strength, and let H , whose components are α, β, γ , represent the magnetic field strength. Furthermore, let D , whose components are D_x, D_y, D_z , represent the dielectric displacement, which is equal to E plus 4π times the electric dipole moment per cubic centimeter in the medium. In general, D is a linear vector function of the E components, such as

$$D_x = \epsilon_{11}X + \epsilon_{12}Y + \epsilon_{13}Z$$

and similarly for D_y and D_z . The nine coefficients ϵ are the **dielectric constants**. In general the direction of D is not the same as the direction of E . However, one may always find at least one set of three mutually perpendicular directions in which the components of D are proportional solely to the corresponding components of E . These are the **principal axes** of the medium. Along these axes

$$D_x = \epsilon_1 X, \quad D_y = \epsilon_2 Y, \quad \text{and} \quad D_z = \epsilon_3 Z$$

where the ϵ 's are the **principal dielectric constants**. In a crystal at least two of them are different, except that in cubic crystals, as in isotropic media in general,

$$\epsilon_1 = \epsilon_2 = \epsilon_3 = \epsilon$$

Maxwell's fundamental equations for insulating media are

$$\begin{aligned} \frac{1}{c} \frac{\partial D_x}{\partial t} &= \frac{\partial \gamma}{\partial y} - \frac{\partial \beta}{\partial z} \\ \frac{1}{c} \frac{\partial D_y}{\partial t} &= \frac{\partial \alpha}{\partial z} - \frac{\partial \gamma}{\partial x} \\ \frac{1}{c} \frac{\partial D_z}{\partial t} &= \frac{\partial \beta}{\partial x} - \frac{\partial \alpha}{\partial y} \end{aligned} \tag{1}$$

$$\begin{aligned}
-\frac{1}{c} \frac{\partial \alpha}{\partial t} &= \frac{\partial Z}{\partial y} - \frac{\partial Y}{\partial z} \\
-\frac{1}{c} \frac{\partial \beta}{\partial t} &= \frac{\partial X}{\partial z} - \frac{\partial Z}{\partial x} \\
-\frac{1}{c} \frac{\partial \gamma}{\partial t} &= \frac{\partial Y}{\partial x} - \frac{\partial X}{\partial y}
\end{aligned} \tag{2}$$

$$\frac{\partial}{\partial x} (\epsilon_1 X) + \frac{\partial}{\partial y} (\epsilon_2 Y) + \frac{\partial}{\partial z} (\epsilon_3 Z) = 0 \tag{3}$$

$$\frac{\partial \alpha}{\partial x} + \frac{\partial \beta}{\partial y} + \frac{\partial \gamma}{\partial z} = 0 \tag{4}$$

The set of equations (1) is a mathematical deduction from Biot-Savart's law for the magnetic field strength about an electric current. Displacement currents in dielectrics are included, but it is assumed that there are no conduction currents. The latter will be considered in the theory of metallic media in Chapter 15. The set (2) is a statement of Faraday's law of induction with the magnetic permeability assumed to be unity, i.e., $B = H$. This simplification is permissible at optical frequencies, even in ferromagnetic media. Equation (3) expresses the continuity of electric flux density in a region containing no free electric charges. Equation (4) is the corresponding expression for the magnetic flux, assuming again that the permeability is unity. Derivations of these formulas are given, for example, in Forsterling's *Lehrbuch der Optik*, pages 49–64, or in Richtmyer and Kennard's *Introduction to Modern Physics*, 3d ed., pages 51–59.

From an atomic viewpoint, the local electric and magnetic field strengths change very rapidly in the vicinity of electrons and nuclei. However, the wave length of light is great compared with atomic dimensions, so that the medium appears practically homogeneous. Thus one need only consider the average values of the fields for most purposes. The fields E and H in the above equations are such averages.

By differentiating equations (1) with respect to t and sub-

stituting the time derivatives of the magnetic vector from (2), one readily finds that

$$\begin{aligned}\frac{\epsilon_1}{c^2} \frac{\partial^2 X}{\partial t^2} &= \nabla^2 X - \frac{\partial}{\partial x} \left(\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} + \frac{\partial Z}{\partial z} \right) \\ \frac{\epsilon_2}{c^2} \frac{\partial^2 Y}{\partial t^2} &= \nabla^2 Y - \frac{\partial}{\partial y} \left(\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} + \frac{\partial Z}{\partial z} \right) \\ \frac{\epsilon_3}{c^2} \frac{\partial^2 Z}{\partial t^2} &= \nabla^2 Z - \frac{\partial}{\partial z} \left(\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} + \frac{\partial Z}{\partial z} \right)\end{aligned}\quad (5)$$

where ∇^2 stands for the operator $\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$. Similar equations may be obtained for α , β , and γ .

First consider an isotropic medium in which $\epsilon_1 = \epsilon_2 = \epsilon_3 = \epsilon$, so that by (3) the equations (5) all reduce to the form

$$\frac{\epsilon}{c^2} \frac{\partial^2 X}{\partial t^2} = \nabla^2 X \quad (6)$$

A particular solution, which is pertinent to this discussion, is

$$X = A \sin \frac{2\pi}{T} \left(t - \frac{lx + my + nz}{V} \right)$$

representing a travelling plane wave, the locus of points in the same phase at a given time being given by

$$lx + my + nz = \text{constant}$$

This is the equation of a plane whose normal has the direction cosines l , m , and n . Substitution into the wave equation (6) shows that

$$\frac{\epsilon}{c^2} = \frac{l^2 + m^2 + n^2}{V^2}$$

or

$$V = \frac{c}{\sqrt{\epsilon}}$$

Thus the index of refraction of the medium $\mu = \sqrt{\epsilon}$.¹ It is only in exceptional cases, however, that one may verify this result by

¹ The symbol μ is used for index of refraction, since n is used for one of the direction cosines.

using the dielectric constant measured at low frequencies, or by using direct currents, and find that its square root is the optical index of refraction. It is well known that μ , as well as ϵ , changes with frequency. Hence it is to be expected that the above relation between them will hold only if both quantities are referred to the same frequency.

That the E waves are necessarily transverse follows from equation (3). Thus, if the amplitude A has the direction cosines L , M , N , then

$$X = LA \sin \frac{2\pi}{T} \left(t - \frac{lx + my + nz}{V} \right)$$

$$Y = MA \sin \frac{2\pi}{T} \left(t - \frac{lx + my + nz}{V} \right)$$

$$Z = NA \sin \frac{2\pi}{T} \left(t - \frac{lx + my + nz}{V} \right)$$

By (3) for isotropic media

$$\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} + \frac{\partial Z}{\partial z} = Ll + Mm + Nn = 0$$

indicating that the cosine of the angle between A and the wave normal is zero, or that the angle between them is 90° . In exactly the same way one may show that the H vibrations are also at right angles to the wave normal. The latter holds for anisotropic as well as isotropic media, whereas the proof for the perpendicularity of the electric field vector does not. However, if the dielectric displacement ϵE is taken as the light vector instead of E , it follows that the waves of D are transverse in any medium.

If one chooses the z axis in the direction of the wave normal, then $n = 1$, and $l = m = 0$. Since the waves are transverse, $Z = 0$ and $\gamma = 0$, and since they are plane waves, $\partial/\partial x = 0$ and $\partial/\partial y = 0$. If one chooses the x axis in the direction of A , then $M = N = 0$, and $L = 1$, so that

$$X = A \sin \frac{2\pi}{T} \left(t - \frac{z}{V} \right)$$

$$Y = 0$$

$$Z = 0$$

Moreover, the Maxwell equations (1) reduce to

$$\frac{\epsilon}{c} \frac{\partial X}{\partial t} = -\frac{\partial \beta}{\partial z}; \quad \frac{\epsilon}{c} \frac{\partial Y}{\partial t} = \frac{\partial \alpha}{\partial z} = 0$$

so that

$$\frac{\epsilon}{c} \frac{2\pi}{T} A \cos \frac{2\pi}{T} \left(t - \frac{z}{V} \right) = -\frac{\partial \beta}{\partial z}$$

Hence

$$\beta = \frac{\epsilon V}{c} A \sin \frac{2\pi}{T} \left(t - \frac{z}{V} \right) = \sqrt{\epsilon} X = \mu X$$

$$\alpha = 0$$

Furthermore, the last equation of (2) indicates that $\partial \gamma / \partial t = 0$; consequently $\gamma = 0$. The constant of integration is taken equal to zero in each case, since possible superposed magnetic fields not associated with the electric field X are of no interest. The resulting electromagnetic wave may be described by the graph shown in Fig. 12.1.

It is seen that the magnetic vector is at right angles to the electric vector and has a magnitude equal to the index of refraction

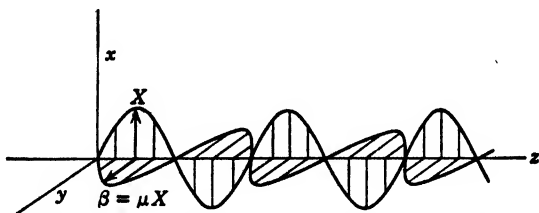


FIG. 12.1. Electric and magnetic vectors of a plane polarized electromagnetic wave.

times the electric vector. The direction of the magnetic vector is such that a 90° rotation of X into β is related to the direction of propagation in the same way as the rotation of a right-handed screw is related to the travel of the screw. The magnitude of the radiant flux per unit area is given by Poynting's vector

$$S = \frac{c}{4\pi} E \times H \text{ ergs/cm}^2 \text{ sec}$$

which in this case becomes

$$S = \frac{c}{4\pi} \mu X^2$$

The flux is pulsating, since it is proportional to $\sin^2 (2\pi/T)t$. Its average value is half its maximum value, or

$$\bar{S} = \frac{c}{8\pi} \mu A^2$$

The intensity of a light beam is thus seen to be proportional to the square of its amplitude, and the constant of proportionality contains the index of refraction as a factor.

12.3 Reflection and Refraction

Let E_p and E_s represent the components of the amplitude of the electric vector of a plane wave incident on an interface, Fig. 12.2, the components being respectively parallel and perpendicular to the plane of incidence. The symbols R_p , R_s , and D_p , D_s , refer to similar components of the reflected and transmitted wave. The phase of each wave is a function of the form

$$\frac{2\pi}{T} \left(t - \frac{lx + my + nz}{V} \right)$$

Let l , m , n represent the direction cosines of the incident wave normal, l_1 , m_1 , n_1 those of the reflected wave normal, and l_2 , m_2 , n_2 those of the refracted wave normal. The existence of boundary conditions at the interface $z = 0$ between the components in the first medium and those in the second requires that the periods be equal and that

$$\frac{l}{V_1} = \frac{l_1}{V_1} = \frac{l_2}{V_2} \quad \text{and} \quad \frac{m}{V_1} = \frac{m_1}{V_1} = \frac{m_2}{V_2}$$

If the x axis is chosen in the plane of incidence, then $m = 0$ and consequently $m_1 = m_2 = 0$, so that the reflected and refracted

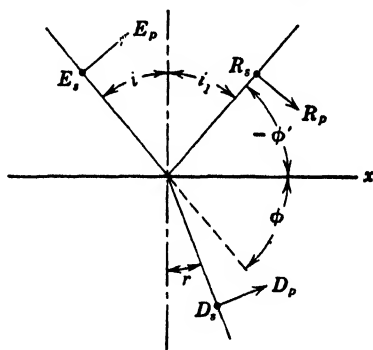


FIG. 12.2. Reflection and refraction of the s and p components of the amplitude of an electromagnetic wave.

wave normals must lie in the plane of incidence. Moreover, $l = l_1$, which means that $\phi = \pm\phi'$ (see Fig. 12.2). The negative sign leads to the law of reflection, $i_1 = i$; the positive sign would give the reflected wave the same direction as the incident wave and is therefore rejected. Furthermore,

$$\frac{l}{V_1} = \frac{l_2}{V_2}$$

leads to

$$\frac{\cos \phi}{\cos \phi_2} = \frac{\sin i}{\sin r} = \frac{V_1}{V_2}$$

which is Snell's law of refraction. The ratio V_1/V_2 is the index of refraction of the second medium with respect to the first. The absolute indices of the media which are referred to vacuo are respectively

$$\mu_1 = \frac{c}{V_1} \quad \text{and} \quad \mu_2 = \frac{c}{V_2}, \quad \text{so that} \quad \mu = \frac{\mu_2}{\mu_1}$$

The standard boundary conditions state that the sums of the tangential components of the electric vectors and of the magnetic vectors are equal in the first and second media at the interface, $z = 0$. In view of the laws of reflection and refraction, the total phase angles are the same on both sides of the interface, so that the sines divide out and one may write for the amplitudes of the electric components

$$E_s + R_s = D_s$$

$$E_p \cos i + R_p \cos i = D_p \cos r$$

and for the tangential components of the magnetic vector

$$\mu_1 E_p - \mu_1 R_p = \mu_2 D_p$$

$$(-\mu_1 E_s + \mu_1 R_s) \cos i = -\mu_2 D_s \cos r$$

The elimination of D_s between the first and the fourth of these relations and the use of $\mu_2/\mu_1 = \sin i/\sin r$ lead to

$$\frac{R_s}{E_s} = -\frac{\sin(i-r)}{\sin(i+r)}$$

Similarly, one finds that

$$\frac{R_p}{E_p} = - \frac{\tan (i - r)}{\tan (i + r)}$$

$$\frac{D_s}{E_s} = \frac{2 \cos i \sin r}{\sin (i + r)}$$

$$\frac{D_p}{E_p} = \frac{2 \cos i \sin r}{\sin (i + r) \cos (i - r)}$$

These are **Fresnel's formulas for reflection and refraction**. When the angles are small, they reduce to

$$\frac{R_s}{E_s} = \frac{R_p}{E_p} = - \frac{i - r}{i + r} = - \frac{\mu - 1}{\mu + 1}$$

and

$$\frac{D_s}{E_s} = \frac{D_p}{E_p} = \frac{2r}{i + r} = \frac{2}{\mu + 1}$$

since $i \doteq \mu r$ for small angles of incidence. The negative sign in the R/E formula indicates a phase reversal if $\mu > 1$. The ratio of the intensity reflected to that incident is given by the square of the amplitude ratio, so that

$$\frac{I_R}{I_E} = \left(\frac{\mu - 1}{\mu + 1} \right)^2$$

Thus, if $\mu = 1.5$, one finds that the reflectivity is 0.04 or 4 per cent. For grazing incidence, $i = 90^\circ$, it is readily seen that the reflectivity is 1.00 or 100 per cent for both s and p components. Between these limits, the reflectivity of the s component steadily rises, while that of the p component drops to zero when $i + r = 90^\circ$ and $\tan (i + r) = \infty$. In this case, since $r = 90^\circ - i$, one finds that

$$\mu = \frac{\sin i}{\sin r} = \frac{\sin i_p}{\cos i_p} = \tan i_p$$

a relation known as Brewster's law. The angle i_p at which the p component is not reflected is called the Brewster angle or the **polarizing angle**, the reflected light being plane polarized even

if the incident light is unpolarized. The curves in Fig. 12.3 show the amplitude and intensity relations for the p and s components of light reflected by crown glass.

The intensity relations for the transmitted components are not obtained directly by squaring the amplitudes, since the media are not the same and the intensity is proportional to μ as well as to E^2 . Moreover, at oblique angles of incidence, the refracted and

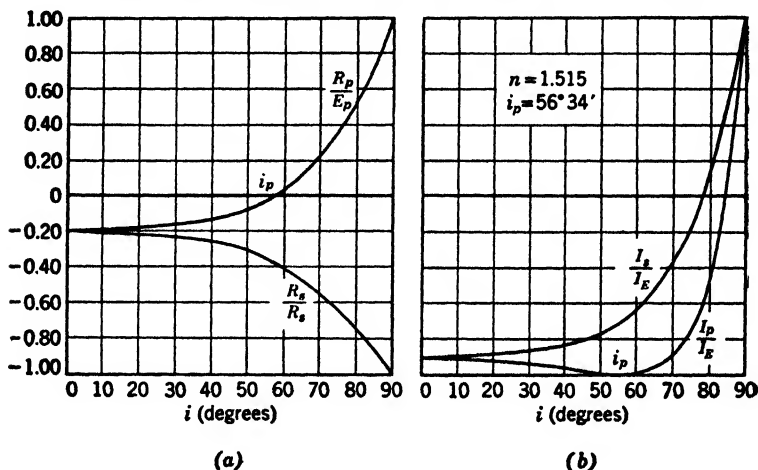


FIG. 12.3. (a) Variation of reflected s and p components of amplitude with the angle of incidence. (b) Variation of reflected s and p components of intensity with angle of incidence for an insulator having an index of refraction of 1.515.

incident beams do not have the same cross-section. When both effects are taken into consideration, it is not difficult to show that the sum of the light flux in the refracted and reflected waves is equal to the light flux in the incident wave, as one would expect. Thus one may find the transmissivity most readily from

$$\frac{I_D}{I_E} = 1 - \left(\frac{R}{E}\right)^2$$

the result being slightly different for the p and s components, indicating partial polarization of the transmitted light.

12.4. Doubly Refracting Media

In 1690 Bartolinus showed that when light was refracted by a calcite crystal there were, in general, two refracted waves. A

direction in which there is no double refraction is called an **optic axis**. Calcite has one such axis and is therefore classed as a **uniaxial** crystal. All crystals which belong to the tetragonal, hexagonal, and rhombohedral systems are uniaxial. Those belonging to the orthorhombic, monoclinic, and triclinic systems are **biaxial**. Crystals of the cubic system are isotropic. The two waves arising from double refraction are plane polarized components of the incident wave. Usually the crystal has a small, but equal, absorption coefficient for both waves. However, some crystals, notably tourmaline and herapathite, absorb one of the components very strongly and are said to be **dichroic**. In a suitable thickness they may be used as polarizing filters for the transmission of one particular component of the incident light. If the incident light is plane polarized, for example by specular reflection, it may be extinguished by suitable orientation of the polarizing filter. This is one of the applications of Polaroid filters, which are made up of tiny herapathite crystals aligned and imbedded in a suitable transparent matrix.

12.5 Crystal Optics

The propagation of light inside an anisotropic medium may be most readily developed by using the electromagnetic equations. A start in this direction has been made in deriving equations (5), page 196. One may choose H , E , or D as the light vector and obtain equations of somewhat different form for the interior of the crystal. However, since light is always observed finally in an isotropic medium, the ultimate result is found to be independent of the choice of the light vector. In order to obtain the formulas first derived by Fresnel for the interior of the crystal, one must choose waves of dielectric displacement D . The components of D are represented by

$$D_x = LA \sin \omega \left(t - \frac{lx + my + nz}{V} \right) = \epsilon_1 X$$

$$D_y = MA \sin \omega \left(t - \frac{lx + my + nz}{V} \right) = \epsilon_2 Y$$

$$D_z = NA \sin \omega \left(t - \frac{lx + my + nz}{V} \right) = \epsilon_3 Z$$

where L , M , and N are the direction cosines of the amplitude A of D . Substituting these equations into the three wave equations (5) gives one the relations:

$$\begin{aligned} -\frac{L}{C^2} &= -\frac{L}{\epsilon_1 V^2} + \frac{l}{V^2} \left(\frac{lL}{\epsilon_1} + \frac{mM}{\epsilon_2} + \frac{nN}{\epsilon_3} \right) \\ -\frac{M}{C^2} &= -\frac{M}{\epsilon_2 V^2} + \frac{m}{V^2} \left(\frac{lL}{\epsilon_1} + \frac{mM}{\epsilon_2} + \frac{nN}{\epsilon_3} \right) \\ -\frac{N}{C^2} &= -\frac{N}{\epsilon_3 V^2} + \frac{n}{V^2} \left(\frac{lL}{\epsilon_1} + \frac{mM}{\epsilon_2} + \frac{nN}{\epsilon_3} \right) \end{aligned}$$

capital C being used to represent the speed of light in free space. These may be simplified by multiplying through by $C^2 V^2$ and letting

$$a^2 \equiv \frac{C^2}{\epsilon_1}, \quad b^2 \equiv \frac{C^2}{\epsilon_2}, \quad c^2 \equiv \frac{C^2}{\epsilon_3}$$

and

$$G^2 \equiv Ll a^2 + Mm b^2 + Nn c^2$$

It is customary to let a represent the greatest, and c the smallest, wave velocity. Then

$$\begin{aligned} L(a^2 - V^2) &= lG^2 \\ M(b^2 - V^2) &= mG^2 \\ N(c^2 - V^2) &= nG^2 \end{aligned} \tag{7}$$

Since

$$lL + mM + nN = 0 \quad \text{and} \quad L^2 + M^2 + N^2 = 1$$

it follows that

$$L^2 a^2 + M^2 b^2 + N^2 c^2 = V^2$$

This is the equation of the **polarization ovaloid**, which shows how the velocity of a wave in a crystalline medium depends on the direction of vibration of the D vector. Its plot in polar coordinates is a surface whose intersection with the x , y , z coordinate planes is shown in Fig. 12.4. For example, in the xy plane, $N = 0$, $L = \cos \theta$, $M = \sin \theta$, one obtains the equation of a Cartesian oval

$$a^2 \cos^2 \theta + b^2 \sin^2 \theta = V^2$$

When $\theta = 0$, one finds that $V = \pm a$, and when $\theta = 90^\circ$, $V = \pm b$, which are two of the principal velocities. Note the single absolute value of V for any given direction of vibration, and vice versa.

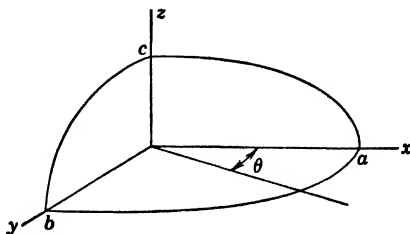


FIG. 12.4. Polarization ovaloid. Dependence of wave velocity on the direction of vibration of the electric vector D .

In terms of the index of refraction $\mu \equiv C/V$ and the principal indices $\alpha \equiv C/a$, $\beta \equiv C/b$, $\gamma \equiv C/c$, the polarization ovaloid may be converted into the index ellipsoid

$$\frac{L^2}{\alpha^2} + \frac{M^2}{\beta^2} + \frac{N^2}{\gamma^2} = \frac{1}{\mu^2}$$

That this is the equation of an ellipsoid with semiaxes α , β , and γ may be shown by considering its intersections with the coordinate planes. For example, in the xy plane

$$\frac{\cos^2 \theta}{\alpha^2} + \frac{\sin^2 \theta}{\beta^2} = \frac{1}{\mu^2}$$

or

$$\frac{(\mu \cos \theta)^2}{\alpha^2} + \frac{(\mu \sin \theta)^2}{\beta^2} = 1$$

which is recognized as the equation of an ellipse when one notes that $x = \mu \cos \theta$ and $y = \mu \sin \theta$ are the rectangular coordinates of any point μ , θ on the curve.

Since by (7)

$$\frac{l}{a^2 - V^2} = \frac{L}{G^2}; \quad \frac{m}{b^2 - V^2} = \frac{M}{G^2}; \quad \frac{n}{c^2 - V^2} = \frac{N}{G^2}$$

it follows from $lL + mM + nN = 0$ that

$$\frac{l^2}{a^2 - V^2} + \frac{m^2}{b^2 - V^2} + \frac{n^2}{c^2 - V^2} = 0 \quad (8)$$

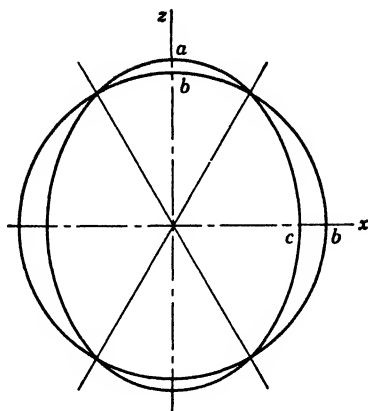


FIG. 12.5. Cross-section of the normal velocity surface, showing the two optic axes of a biaxial medium.

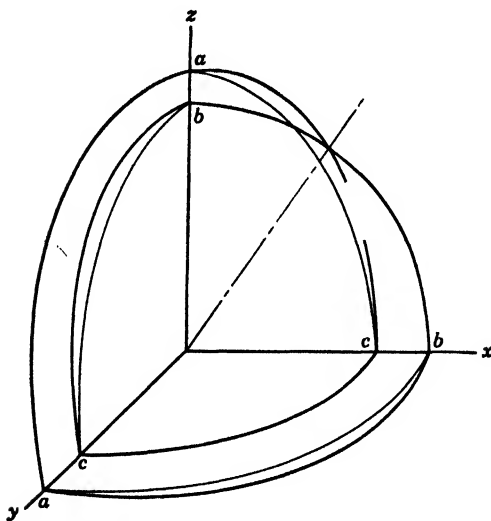


FIG. 12.6. The heavy outline traces one octant of the normal velocity surface. The lighter curves are ellipses which indicate the shape of an octant of the ray velocity surface, Art. 12.6, the circles connecting the ends of the ellipses being common to both ray velocity and wave velocity surfaces.

This is the equation of Fresnel's **wave velocity surface** which, as will be seen later, is less ambiguously called the **normal velocity surface**. Clearing of fractions, one finds that

$$l^2(b^2 - V^2)(c^2 - V^2) + m^2(a^2 - V^2)(c^2 - V^2) + n^2(a^2 - V^2)(b^2 - V^2) = 0$$

The intersections of this double surface with the coordinate planes may be found by setting $l = 0$, $m = 0$, and $n = 0$ in turn. Thus, for the xz plane, $m = 0$, $l = \cos \theta$, $n = \sin \theta$. The two remaining factors give $V^2 = b^2$, a circle, and $V^2 = c^2 \cos^2 \theta + a^2 \sin^2 \theta$, a Cartesian oval whose semiaxes are c and a respectively, as shown in Fig. 12.5. The two curves intersect at four points which define two directions of single wave velocity. These directions are called **bi-normals** or **optic axes**. In a similar manner one may find the intersections of the wave velocity surface with the other two coordinate planes. Each consists of a circle and an oval, but these intersect only in the xz plane. One octant of the surface is shown in Fig. 12.6. The difference between the velocities is greatly exaggerated.

12.6 Ray Velocity Surface

The direction of the light ray in an anisotropic medium is given by the direction of the energy flow or Poynting's vector $S = C/4\pi E \times H$. The rays are therefore at right angles to E and H .

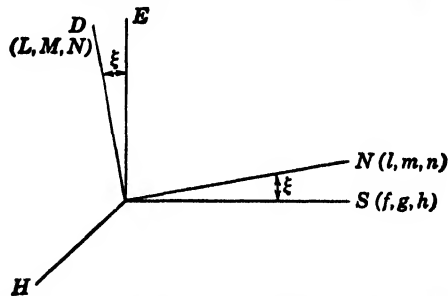


FIG. 12.7. Angular relation between the electromagnetic vectors D , E , H , and the ray direction S and the wave normal direction N .

Since the direction of D is not the same as that of E except along the principal axes, the light rays do not coincide with the wave normals in general. Their space relations are illustrated in Fig. 12.7, where ξ indicates the angle between D and E . The direction

cosines of the ray S will be indicated by f , g , and h , and the velocity in the direction of S will be denoted by U . It can be shown (for example, consult Drude's *Theory of Optics*) that the relation between U and f , g , and h is given by the equation:

$$\frac{f^2}{\frac{1}{a^2} - \frac{1}{U^2}} + \frac{g^2}{\frac{1}{b^2} - \frac{1}{U^2}} + \frac{h^2}{\frac{1}{c^2} - \frac{1}{U^2}} = 0$$

This is the equation of a surface of two sheets, which is called the **ray velocity surface** or, sometimes, the **wave surface**. The latter name originates from the fact that a Huygens wavelet spreading from any point inside the crystal takes the shape of this surface.

12.7 Index Surface

Still another useful relation is that between the index of refraction μ and the direction of the wave normal l , m , n . It may readily be obtained from the normal velocity equation by elimination of a , b , c , and V , in favor of the quantities α , β , γ , and μ which are respectively proportional to their reciprocals. The resulting double surface is called the **index surface** because it describes the variation of the two indices of refraction corresponding to any given direction of the wave normal. In general shape, this surface is like the ray velocity surface, the intersections with the coordinate planes being circles and ellipses.

12.8 Velocities inside Anisotropic Media

The two possible wave velocities which correspond to a given direction of the wave normal l , m , n , may be obtained by solving the biquadratic equation, page 207, for V^2 . The result is conveniently expressed in terms of the quantities defined by the following set of equations

$$I = l^2(b^2 + c^2); \quad J = m^2(c^2 + a^2); \quad K = n^2(a^2 + b^2)$$

$$R = l^2(b^2 - c^2); \quad S = m^2(c^2 - a^2); \quad T = n^2(a^2 - b^2)$$

The two values of V^2 can then be written in the form

$$2V^2 = I + J + K \pm \sqrt{(R + S - T)^2 - 4RS}$$

Equations from which the wave velocities are sometimes more readily calculated are obtained if one expresses the above result in terms of the angles g_1 and g_2 , which the wave normal makes with the two optic axes. The positive directions of these axes are chosen as those which make an acute angle with the positive direction of the z axis. Their direction cosines are:

$$l_1 = + \left(\frac{a^2 - b^2}{a^2 - c^2} \right)^{1/2}; \quad m_1 = 0; \quad n_1 = + \left(\frac{b^2 - c^2}{a^2 - c^2} \right)^{1/2}$$

and

$$l_2 = - \left(\frac{a^2 - b^2}{a^2 - c^2} \right)^{1/2}; \quad m_2 = 0; \quad n_2 = + \left(\frac{b^2 - c^2}{a^2 - c^2} \right)^{1/2}$$

The angles between the wave normal and these axes are then given by the usual equations

$$\cos g_1 = ll_1 + mm_1 + nn_1$$

and

$$\cos g_2 = ll_2 + mm_2 + nn_2$$

In terms of these two angles the roots of Fresnel's wave velocity equation (8) are

$$2V_1^2 = (a^2 + c^2) + (a^2 - c^2) \cos (g_1 - g_2)$$

$$2V_2^2 = (a^2 + c^2) + (a^2 - c^2) \cos (g_1 + g_2)$$

By replacing the velocities by the reciprocals of the corresponding indices of refraction, one obtains equations for the two indices of refraction associated with a given wave normal.

To each of the two values of V there will correspond only one direction of the light vector D . The two sets of direction cosines, one for each V , are given by the ratios

$$L_1:M_1:N_1 = \frac{l}{a^2 - V_1^2} : \frac{m}{b^2 - V_1^2} : \frac{n}{c^2 - V_1^2}$$

$$L_2:M_2:N_2 = \frac{l}{a^2 - V_2^2} : \frac{m}{b^2 - V_2^2} : \frac{n}{c^2 - V_2^2}$$

The exact values of L , M , N , are found from these ratios with the aid of the requirement that the sum of the squares of the direction cosines must equal unity. It can be readily proved that the two vibration directions are at right angles to each other. In

uniaxial substances one of them is in the plane of the wave normal and the optic axis, whereas the other is at right angles to this so-called **principal plane**. In biaxial crystals or substances, the wave normal defines a plane with each optic axis. In this case the *D* vibrations lie in the two bisecting planes. These planes also contain the rays which correspond to the respective vibrations. It should be noted that these statements and equations apply only to the propagation of light wholly inside the anisotropic medium. Thus the two values of *V* and their associated vibration directions and light rays are, in general, not the two observed when light passes from air into the medium, since in the general case of double refraction by a crystal there are two waves of *different* *l*, *m*, and *n*. In the above discussion it is assumed that the same wave normal applies to both waves. In the passage of light from one medium to another, this is true only if the angle of incidence is zero.

It was noted before that the directions of energy flux, the **light rays**, are in general not the same as the wave normals. It is therefore necessary to distinguish between the velocity of propagation *U* of the light in the direction of the ray and the velocity of the wave *V*, the relation being $U = V / \cos \xi$, where ξ is the angle between the light ray and the wave normal, Fig. 12.7. The ray velocity associated with any given wave velocity is accordingly always greater than the latter; and still, strangely enough, when one plots *U* on the same polar diagram as *V*, the former surface lies within, and at most touches, the latter. The relation between the two surfaces is shown in Fig. 12.6.

12.9 Double Refraction

The ray velocity surface is of particular importance in that it gives a picture of the manner in which light spreads out in all directions from a point inside the crystal. Thus it gives the form of the Huygens wavelets in an anisotropic medium and may be used in the determination of the reflected and refracted rays and waves resulting from the passage of light through an interface between two media of which one or both may be anisotropic.

The application of Huygens' principle to the general case of oblique refraction at the surface of a biaxial substance is illustrated in Fig. 12.8. The double wavelet from the point *C* spreads out as shown during the time that the incident wave is travelling from *B* to *D*. The incident wave is assumed to be perpendicular to the

paper, or plane of incidence. The two refracted waves are also found to be perpendicular to the plane of incidence, and are respectively tangent to the two sheets of the wave surface. The corresponding wave normals N_1 and N_2 consequently lie in the plane of incidence. The two rays are found by drawing lines from the point C to the points of tangency of the two refracted waves and the wave surface. In general, these rays will not lie in the plane of incidence; however, it is impossible to show this in a two-dimensional figure. If the refracted wave is travelling in the direction of an optic axis in the crystal, it will touch the outer wave surface at an infinite number of points lying on a circle around one of the depressions in the wave surface. There will then be an infinite number of refracted rays lying on a cone, and because of this the phenomenon is called **internal conical refraction**.

In many studies of double refraction it is customary to have the light strike the surface perpendicularly. This always results in two refracted waves whose normals coincide with each other and are also normal to the surface. The rays, however, may be at various angles with the normal.

If there exist any special relations among the principal velocities a , b , and c , this fact simplifies the problem. If any two of the principal velocities are equal, the wave surface consists of a sphere and a spheroid tangent to each other at two points which define the direction of the single optic axis. The spheroid is prolate or oblate, depending on whether the crystal is positively or negatively uniaxial.

Two applications of Huygens' principle to refraction at the interface between an isotropic and a negatively uniaxial material are shown in Fig. 12.9. The optic axis is assumed to lie in the plane of incidence. The directions of the rays are given by the lines from C to the points of tangency, while the wave normals N_e and N_o are

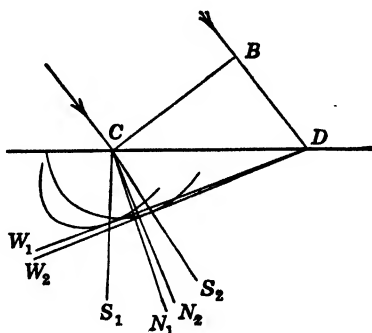


FIG. 12.8. Application of Huygens' principle to the general case of oblique refraction at the surface of a biaxial crystal.

by definition the perpendiculars to the two waves. The ray and wave normal belonging to the wave determined by the spherical wavelets are coincident. For this wave the same laws of refraction hold as for isotropic media, and for this reason the corresponding ray is called the **ordinary ray**, although it is plane polarized. The other, **extraordinary ray**, is not in the direction of the corresponding wave normal and does not obey the ordinary laws of refraction. This ray is also plane polarized. The D or E vibrations associated with the ordinary ray are normal to the principal plane

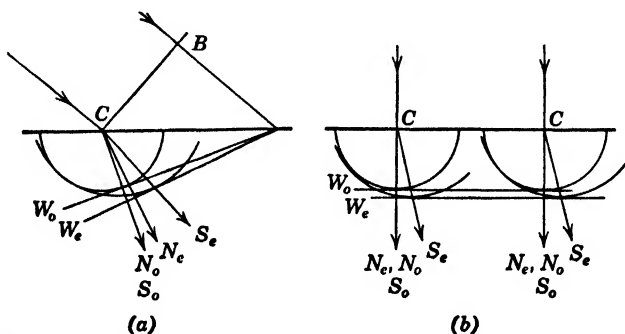


FIG. 12.9. (a) Oblique refraction at the surface of a uniaxial crystal. (b) Refraction resulting from normal incidence on a uniaxial crystal.

defined by the ordinary ray and the optic axis. The D vibrations corresponding to the extraordinary wave normal lie in the principal plane formed by this wave normal and the optic axis. When the optic axis lies in the plane of incidence, the extraordinary and ordinary vibration directions are exactly at right angles to each other. This is true also whenever the incidence is normal, as is illustrated in Fig. 12.9(b). The two directions of vibration are called the **principal vibration directions** of the plate. The one for the extraordinary ray is given by the projection of the optic axis on the surface of the uniaxial plate. Any incident plane polarized wave is resolved into components vibrating in this direction and at right angles to it.

Incident plane polarized light may be obtained in a variety of ways. One convenient method is by the use of polarizing filters in sheet form, such as those marketed under the name of Polaroid. These plates contain a layer of uniformly oriented microscopic crystals of herapathite between glass plates or celluloid sheets.

Herapathite is an iodosulphate of quinine which is dichroic, meaning that it not only resolves an incident vibration into two orthogonal components, but absorbs one of them much more strongly than the other. In a thickness great enough to absorb one component almost completely, the other transmitted component is also reduced appreciably. The degree of polarization falls off at both ends of the visible spectrum, making this a somewhat inefficient, though conveniently compact, polarizer.

A method by which one may obtain a polarized beam of large cross-section with the least expense is by the use of a reflecting

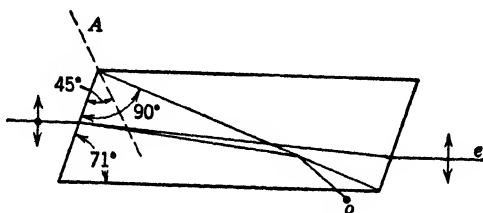


FIG. 12.10. Passage of rays through a Nicol prism. The direction of the optic axis is indicated by A. The ordinary ray *o* is totally reflected, leaving the *e* ray to be transmitted.

glass plate at the polarizing angle. This method employs the discovery made by Malus and Brewster that the light reflected by any isotropic insulator is plane polarized when the tangent of the angle of incidence is equal to the index of refraction of the substance. The reflecting power of a glass plate at this angle, which is about 57° for ordinary glass, is only 14 per cent for the component reflected. Surface films and other imperfections diminish the degree of polarization. The reflected electric vibrations are predominantly normal to the plane of incidence.

The most efficient method for producing plane polarized light is still the **Nicol prism** or one of its modifications. The Nicol prism is made of calcite, the ends of which are ground to make the angles shown in Fig. 12.10. This prism is then cut diagonally at such an angle that, when cemented together by Canada balsam, the ordinary ray is totally reflected at the balsam layer and absorbed in the blackened mounting of the prism. Only the extraordinary ray, vibrating in the principal plane which it makes with the optic axis, is transmitted. This plane of vibration is parallel to the short diagonal of the rhombic cross-section of the prism. In order that

the conditions for total reflection be satisfied, the direction of the incident light must not deviate too much from the direction of the long axis of the Nicol prism. This limits the aperture to about 29° . Of the other similar forms of polarizing prism, the most common is that of Glan-Thompson, which has perpendicular ends and a square cross-section. It is somewhat more efficient as a polarizer but has a smaller field of view.

In some studies it is necessary to produce wide beams of polarized light and to have the rays incident normally on the specimen. If a Nicol prism is used alone as a polarizer, this would require one of

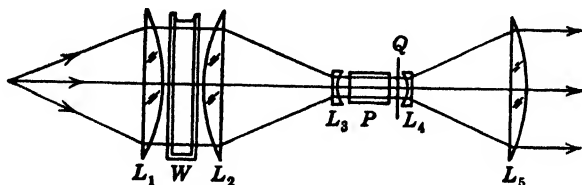


FIG. 12.11. Polarizing unit for obtaining a wide beam of polarized light from a small polarizer P . The quarter-wave plate Q is inserted when circularly polarized light is desired, Art. 12.13.

very large cross-section, the cost of which would be prohibitive. Hence an arrangement of lenses similar to that shown in Fig. 12.11 is employed to accomplish the desired result with the use of a small polarizing prism. The diverging lenses L_3 and L_4 keep the rays which pass through the relatively small Nicol prism well within its aperture. In fact, the rays are practically parallel and thus assure that the vibration direction is the same over the entire field of view. If they were not parallel, one can see that, as a consequence of the different principal planes for each different direction of wave propagation, the vibrations associated with each ray would be in a somewhat different direction.

Since the principal planes in any application of crystal optics vary with the direction of the light, it is often more convenient to refer to the principal section, which is fixed by the orientation of the surface with respect to the optic axis. A **principal section** is a plane containing the optic axis and the normal to the crystal face or crystal cut. A principal section is also a principal plane when the light is incident perpendicularly so that the wave normal lies in the principal section. Since the angle of incidence is often

zero, the two terms are sometimes used interchangeably, although a principal section is not always a principal plane.

12.10 Plane Polariscopes

A plane polariscopes consists of two polarizing filters, which may be Nicol prisms, mounted on a common axis. These are called the **polarizer** and **analyzer**, respectively. When one of them is rotated with respect to the other, the transmitted light intensity varies in accordance with Malus' cosine-squared law. If one sets the vibration planes of the polarizer and analyzer at 90° to each other, or at some odd multiple of 90° , the light is completely extinguished. If one now inserts a plate of anisotropic material between the polarizer and analyzer, some of the original light intensity is restored unless the light happens to be travelling along an optic axis or unless either one of the principal vibration directions of the plate coincides with the principal plane of the polarizer or analyzer. There will be a change in the color of the transmitted light when monochromatic light is not used. Since the color is due to a superposition of interference effects in the different wave lengths, it is logical to consider monochromatic light at first.

When polarized light of wave length λ enters the doubly refracting plate, its two components in the principal vibration directions of the plate are propagated with different speeds so that they emerge with the phase difference

$$\Delta = \frac{2\pi}{\lambda} (\mu_1 - \mu_2)d$$

Their amplitudes will be respectively $a' \cos \phi$ and $a'' \sin \phi$, where a' and a'' are the products of the incident amplitude and the amplitude transmission coefficients of the plate for the two vibrations. Reflection losses at the two surfaces are included in these factors. Since for normal incidence on transparent media, the losses due to both causes are very nearly equal, we can put $a' = a'' = a$ for most practical purposes. The resultant of two perpendicularly vibrating polarized and coherent waves with a phase difference Δ is in general an elliptical vibration, Art. 12.11. The intensity of such a combination is independent of ϕ and Δ . However, when this light is passed through the analyzer, only the coplanar components of $a \cos \phi$ and $a \sin \phi$ which lie in the principal section of the analyzer will be transmitted. These retain the phase dif-

ference Δ , since they both travel with the same speed over the same distances except in the doubly refracting plate. The amplitudes of the transmitted wave components are obtained by a vector diagram of the type shown in Fig. 12.12, where A represents the principal plane of the analyzer, and P , that of the polarizer; while x and y denote the principal vibration directions of the plate being examined. If the principal planes of the polarizing filters are at

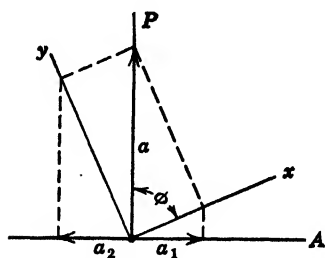


FIG. 12.12. Amplitude components transmitted by a plane polariscope with the analyzer and polarizer crossed.

right angles to each other, as shown in Fig. 12.12, the wave components transmitted by the analyzer have the amplitudes:

$$a_1 = a \cos \phi \sin \phi;$$

$$a_2 = -a \sin \phi \cos \phi$$

These are equal numerically, the minus sign indicating a reversal in phase which may be either applied to the amplitude, as above, or regarded as a change in Δ by the amount π . In either case, the resultant of the two waves, Art. 10.2, is a plane wave whose amplitude is given by

$$A^2 = a^2(2 \cos^2 \phi \sin^2 \phi - 2 \cos^2 \phi \sin^2 \phi \cos \Delta) = a^2 \sin^2 2\phi \sin^2 \frac{\Delta}{2}$$

If the polarizing filters are parallel, it is easy to show that the equation for A_{\parallel}^2 becomes

$$A_{\parallel}^2 = a^2 \left(1 - \sin^2 2\phi \sin^2 \frac{\Delta}{2} \right)$$

In many investigations it is customary to use crossed polarizers, to which the *first* equation applies. It shows that the intensity of the light transmitted by a plane polariscope depends on the orientation of the specimen and on the phase difference which it introduces. Note in particular that, regardless of the value of Δ , A^2 is zero when $\phi = 0, \pi/2, \pi, 3\pi/2$; that is to say, the intensity is zero when either of the vibration directions of the plate coincides with the principal plane of the polarizer or of the analyzer, these being at right angles. By looking for this condition while rotating the analyzer and polarizer together so as to preserve the 90° setting,

one may locate the principal vibration directions of the material being examined in the polariscope. This procedure is used in photoelastic studies to find the orientation of the stress ellipsoid, since its axes coincide with the principal vibration directions. The lines of zero intensity observed with any fixed setting of the polarizing filters or prisms are called **isoclinics**.

For any given value of ϕ , the intensity depends on Δ , as shown in Fig. 12.13. Complete darkness occurs when $\Delta = m2\pi$, where m is any integer and is called the order of interference. If the

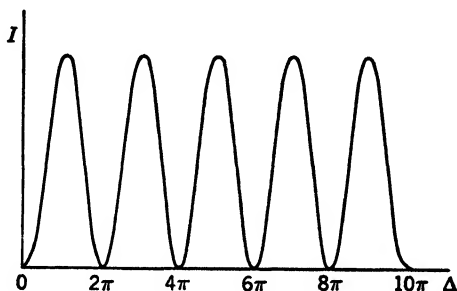


FIG. 12.13. Dependence of intensity on the phase difference of two waves having equal amplitudes and an initial phase difference of π , for example, a_1 and a_2 in Fig. 12.12.

value of Δ varies over the surface of the plate, there will be dark bands or interference fringes in the field of view. Along any one band, the value of Δ is constant for any given wave length, and therefore $(\mu_1 - \mu_2)d$ is constant. If the thickness d is uniform, these fringes are loci of constant **birefringence**, $(\mu_1 - \mu_2)$. In photoelastic applications this quantity is proportional to the difference between the principal stresses in the plane of the plate. Accordingly, the stress difference is uniform along any one fringe and varies from fringe to fringe by a constant increment which depends on the stress-optical constant of the material employed.

It has been assumed above that monochromatic light is being used. If white light is used, it is due primarily to the variation in the magnitude of the wave length itself, and secondarily to the dispersion in the birefringence, that the value of Δ differs for the various spectral colors. Thus in the transmitted light some wave lengths will be in the same phase and some in opposite phase, producing a considerable variation in intensity throughout the

spectrum. This **selective interference** will cause the transmitted light to be colored except where the birefringence is zero. Here the phase difference is zero for all wave lengths, and the band is achromatic, being perfectly black when the polarizing filters are crossed, and white when they are parallel. The locus of points of the same color are called **isochromatic lines**. Because of the maximum sensitivity of the eye to light of a yellow-green color, the complementary purple locus is quite well defined in several orders. The colored bands fade out as Δ becomes large, and in any case they are not as sharp as the interference fringes in monochromatic light, and hence do not lend themselves readily to quantitative work. The principal value of white light is that it facilitates the location of the fringe of zero order which shows where the birefringence is zero.

12.11 Elliptically and Circularly Polarized Light

The light vector in polarized light may not vary in a simple harmonic manner in one plane, as in plane polarized light, but may rotate with or without change in magnitude. When this is the case, the light is said to be **elliptically** or **circularly polarized**. Such light may be considered to be the resultant of two simple harmonic vibrations at right angles to each other with some phase difference between them. This is how it is usually produced. When plane polarized light passes perpendicularly through any anisotropic medium, two orthogonally polarized waves emerge, as was pointed out above. These two waves may be represented by the equations:

$$x = a_1 \sin \left[\frac{2\pi}{T} \left(t - \frac{z}{V_1} \right) + \epsilon \right]$$

$$y = a_2 \sin \left[\frac{2\pi}{T} \left(t - \frac{z}{V_2} \right) + \epsilon \right]$$

where $a_1 = a \cos \phi$ and $a_2 = a \sin \phi$ are the two amplitudes, and x and y denote the components of the light vector in the corresponding principal vibration directions of the plate. If z is measured from the front surface of the plate, its value at emergence will be equal to the thickness of the plate, and, because of the difference between V_1 and V_2 , there will be a phase difference given by the usual equation. After passing through the plate the phase *difference* remains constant, although the phase itself is a function of

z and t . At any given point, z , there are two simple harmonic vibrations of the form:

$$x = a_1 \sin \omega t$$

$$y = a_2 \sin (\omega t + \Delta)$$

The resultant vibration of which x and y are components is, in general, no longer plane polarized, since the instantaneous values of x and y are related by the equation:

$$\frac{x^2}{a_1^2} - \frac{2xy}{a_1 a_2} \cos \Delta + \frac{y^2}{a_2^2} = \sin^2 \Delta$$

which is readily obtained from the above equations by eliminating the time t . This equation shows that the end of the light vector, the point x, y , describes an ellipse. The presence of the xy product term shows that, in general, the ellipse is inclined with respect to the x and y directions. There are several important special cases. When $\Delta = 0, 2\pi, 4\pi, \dots, m2\pi$, the equation reduces to that of a straight line with the slope a_2/a_1 , which is the same as that of the original vibration. On the other hand, when $\Delta = \pi, 3\pi, 5\pi, \dots, (m + \frac{1}{2})2\pi$, the slope is $-a_2/a_1$, and the vibration direction makes an angle equal to $2 \tan^{-1} a_2/a_1$ with its original direction. A plate of such a thickness that $\Delta = \pi$ is called a **half-wave plate** and is frequently used to rotate a plane of polarization through any desired angle for a particular color. Since the magnitudes of a_1 and a_2 are the sine and cosine components of the incident light vector, their relative values can be varied at will by rotating the half-wave plate.

When $\Delta = \pi/2, 3\pi/2, 5\pi/2, \dots, (m + \frac{1}{2})\pi$, the vibration equation reduces to that of an ellipse

$$\frac{x^2}{a_1^2} + \frac{y^2}{a_2^2} = 1$$

the axes of which coincide with the x and y directions and have the values $2a_1$ and $2a_2$. A plate for which $\Delta = \pi/2$ is called a **quarter-wave plate**. As before, the quantities a_1 and a_2 can be given any relative value by rotating the plate, that is, by changing the azimuth angle ϕ . In this way the ellipse can be given an eccentricity of any desired value. In particular, if ϕ is equal to 45° , $a_1 = a_2$, and the above equation of the form of the vibration re-

duces to that of a circle. The electric vector at any point simply rotates about the wave normal without change in its magnitude. This kind of light is said to be **circularly polarized**.

In order to determine the existence and kind of polarization of a beam of light, one may use a Nicol prism or other polarizing filter, and a quarter-wave plate. The beam is first tested by passing it through only the Nicol prism which is being slowly rotated. If the light can be extinguished, it must be plane polarized. If its intensity does not change, it is either unpolarized or circularly polarized. In the latter case, a quarter-wave plate inserted ahead of the Nicol prism in any azimuth will change it to plane polarized light, which can then be extinguished by the Nicol prism.

If the intensity varies as the Nicol prism is rotated, without interposing a quarter-wave plate, the light is either partially plane polarized or elliptically polarized. In the latter case, one may find, with a quarter-wave plate placed ahead of the Nicol prism, two azimuths in which the elliptically polarized light is changed to plane polarized light. These two azimuths are 90° apart and correspond to the coincidence of the "fast" axis of the quarter-wave plate with one or the other of the axes of the ellipse. The tangent of half the angle between the corresponding extinction positions of the Nicol prism gives the ratio of the axes of the ellipse.

12.12 Babinet Compensator

The Babinet compensator is useful in measuring phase differences caused by double refraction. It is made of two wedges of

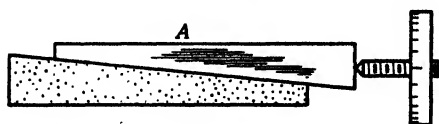


Fig. 12.14. Section through the wedges of a Babinet compensator.

quartz with their optic axes at right angles to each other and in the directions of the length and width of the two respective wedges, Fig. 12.14. One of the wedges can be displaced by means of a calibrated screw. At some point such as *A*, where the thicknesses of the wedges are equal, the phase difference will be zero. It will increase linearly in both directions from *A*. This will result in the

transmission of any desired ellipticity, including circularly and plane polarized light, at various points along the wedge. At regularly spaced intervals, the phase difference changes by an increment of 2π , which means that the original vibration is transmitted. Between crossed Nicols, this light is extinguished, so that one sees a set of equally spaced dark bands. The screw rotation required to displace the bands by one interval, that is, change the phase difference by 2π , gives the required information for translating a fringe shift of any magnitude into the corresponding phase difference by direct proportion. If the axes of the compensator coincide with the axes of the incident elliptically polarized light, the phase difference observed will be just $\pi/2$. This fact is often applied to the finding of the orientation of the axes of elliptically polarized light. In the general measurement of phase differences, the axes of the compensator must be aligned with the vibration components between which the phase difference is desired. The screw rotation required to displace the zero order band to its original position at A gives a measure of the phase difference.

In a Soliel-Babinet compensator, the two wedges have their principal sections parallel to each other, and they are supplemented by a parallel plate having its principal section at right angles to the principal section of the wedges and having a thickness equal to that of the combined wedges in their zero position. The effect is then equivalent to that of a doubly refracting plate of uniform but variable thickness. Consequently no bands are seen, the intensity being uniform but undergoing periodic variations as one slides one wedge over the other. In white light, naturally, one obtains colors which run through the various orders of the Newton color scale.

The Babinet compensator measures only the *difference* between the phases of the two components transmitted by any anisotropic plate. From this one may obtain the relative retardation $(\mu_1 - \mu_2)d$, or the birefringence when the retardation is divided by the thickness of the plate.

12.13 Interference in Circularly Polarized Light

It is often of advantage to examine anisotropic materials with the aid of a circularly polarizing polariscope. A quarter-wave plate is placed behind the usual plane polarizer, with its fast axis at $\pm 45^\circ$ with the direction of the incident plane polarized light.

A clockwise rotation of the fast axis, as in Fig. 12.15, gives right-hand circularly polarized light, and the combination of Nicol prism followed by the quarter-wave plate is called a **right-hand circular polarizer**. In this case, the rotation of the electric vector is related to the direction of propagation in the same sense as the rotation of a right-hand screw is related to its translation. The observer sees it as a counterclockwise rotation. A similar combination with plate and Nicol in the reverse order is used at

the other end as a right-hand circular analyzer. The circularly polarized light transmitted by the polarizer may be represented by the equations of its components

$$x = a \cos \omega t = a \sin \left(\omega t + \frac{\pi}{2} \right)$$

$$y = a \sin \omega t$$

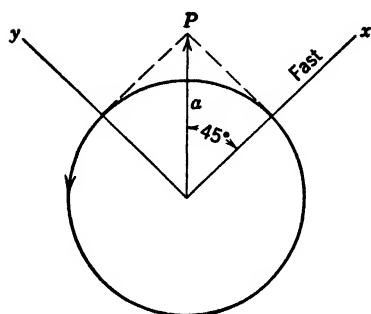


FIG. 12.15. Production of right-hand circularly polarized light. The fast direction of the quarter-wave plate is 45° clockwise of the principal section P of the polarizer.

Suppose that these waves strike a right-hand circular analyzer which has been rotated through 90° , so that all corresponding vibration directions are at right angles to those of the polarizer.

The slow axis of the analyzing quarter-wave plate will then be in the x direction, so that the components it transmits will be

$$x = a' \sin \left(\omega t' + \frac{\pi}{2} - \frac{\pi}{2} \right)$$

$$y = a' \sin \omega t'$$

Their resultant is in the original direction P , as can be seen by examining Fig. 12.15. Therefore light is not transmitted by the analyzing Nicol prism which is at right angles to P .

If now a doubly refracting plate is placed between the polarizer and analyzer of such a circular polariscope, it will transmit two waves vibrating in the directions of its principal axes which make some angle ϕ with P . To find these wave components, one resolves the incident circularly polarized light into components along the vibration directions of the plate. Since circularly polarized

light has the same components for any mutually perpendicular directions, one may use the same equations as given at first, except that now the x and y refer to the principal directions of vibrations transmitted by the substance to be examined. These components are transmitted with a slight reduction in amplitude and with the introduction of a relative phase retardation Δ , so that they may be represented by

$$x' = a'' \sin \left(\omega t'' + \frac{\pi}{2} + \frac{\Delta}{2} + \frac{\Delta}{2} \right)$$

$$y' = a'' \sin \left(\omega t'' + \frac{\Delta}{2} - \frac{\Delta}{2} \right)$$

The double prime on t takes care of the phase retardation in the plate, and those on a take care of the reflection and absorption losses, which are usually small and equal for both components. When light is incident normally and the substance is not dichroic, the primes may conveniently be dropped. The phase difference Δ has been split into halves in order to make it easier to show that x' and y' represent a right-hand and a left-hand circularly polarized wave travelling together. To do this, these equations are expanded by the trigonometric formulas for the sine of the sum and difference of two angles, as follows:

$$x' = a \sin \left(\omega t + \frac{\pi}{2} + \frac{\Delta}{2} \right) \cos \frac{\Delta}{2} + a \cos \left(\omega t + \frac{\pi}{2} + \frac{\Delta}{2} \right) \sin \frac{\Delta}{2}$$

$$y' = a \sin \left(\omega t + \frac{\Delta}{2} \right) \cos \frac{\Delta}{2} - a \cos \left(\omega t + \frac{\Delta}{2} \right) \sin \frac{\Delta}{2}$$

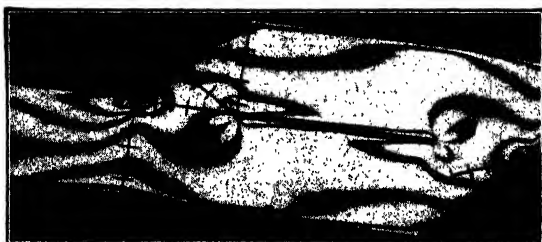
The first terms of x' and y' together represent waves of equal amplitude and a phase difference of $\pi/2$ and therefore make up a right-hand circularly polarized wave having an amplitude equal to $a \cos \Delta/2$. Similarly, the remaining terms represent a left-hand circularly polarized wave having the amplitude $a \sin \Delta/2$. Since a right-hand circular analyzer set for extinction will transmit only the latter, the transmitted light has an intensity which is proportional to $\sin^2 \Delta/2$. It is important to note that the intensity is the same regardless of the orientation of the principal axes of the plate, being independent of ϕ , but otherwise dependent on the phase difference Δ in exactly the same way as the light transmitted by a



(a)



(b)



(c)



(d)

FIG. 12.16. Photoelasticity. Each band is an isochromatic line which traces the locus of constant difference between the principal stresses. The stress difference changes by equal increments from one band to the next. The figures (a), (b), (c), (d) show the effect of a progressive increase in stress because of tension across a model of a lap weld. Circularly polarized light of 5461 \AA was employed. (Photographs by Professor J. J. Ryan, University of Minnesota.)

plane polariscope with its Nicols crossed. In the latter case, the interference due to variation in Δ is obscured when the value of ϕ approaches 0° or 90° , Art. 12.10.

The circular polariscope is frequently utilized in the photo-elastic study of stress distribution in transparent models, Fig. 12.16. A given difference between the principal stress components in the plane of the model produces the same interference effect regardless of the orientation of the stresses when circularly polarized light is used. In plane polarized light, on the other hand, the interference bands disappear when the stresses are either parallel or perpendicular to the principal plane of the polarizer. This fact is often used to determine the *direction* of the stresses, while circularly polarized light is used to determine the magnitude of their *difference*. Their *sum* is found by measuring the changes in thickness of the stressed plate by means of an ordinary interferometer. From the sum and difference, the magnitudes of the stresses may be found.

PROBLEMS

1. The index of refraction of Canada balsam is 1.530, and the index of calcite is 1.658 for the ordinary ray and 1.486 for the extraordinary ray. Which of these rays may be totally reflected in the calcite by a Canada balsam film and under what conditions?

2. The polarizer and analyzer of a polariscope have their principal planes parallel, and a crystal plate is inserted between them with its principal section at 30° with that of the polarizer. What is the amplitude ratio of the components passing through (a) the crystal plate? (b) the analyzer?

3. Quartz has a refractive index of 1.544 for the ordinary ray and 1.553 for the extraordinary ray, when measured with sodium light. What thicknesses of quartz between a crossed polarizer and analyzer will produce annulment of this light, the quartz being cut parallel to its optic axis?

4. (a) In what general respects is the explanation of the colors of anisotropic media in a polariscope with crossed polarizers similar to the explanation of the colors of thin films in reflected light? (b) How do they differ?

5. Calcite has a refractive index of 1.658 for the ordinary ray and 1.486 for the extraordinary ray. What wave lengths in the visible range will be annulled when a thickness of 0.01 mm cut parallel to the optic axis is placed between a polarizer and analyzer whose principal sections are (a) parallel? (b) perpendicular?

6. A crystal plate 0.1 mm thick having principal indices of 1.5590 and 1.5567 is placed in a polariscope. (a) What is the difference in optical path for the transmitted components? (b) What is the difference in phase in degrees for light whose wave length is 5500 \AA ?

7. A Rochon prism is made of calcite, each oppositely directed half prism having a refracting angle of 30° with opposite angles of 60° and 90° . What is

the angle between the transmitted beams if the incidence is normal? In a **Rochon prism** the optic axis in one of the half prisms is perpendicular to its face and to its refracting edge, whereas in the other half it is parallel to the face and to the refracting edge.

8. Plane polarized sodium light is incident normally on a birefringent plate of 0.01-mm thickness, whose principal indices are 1.5590 and 1.5567. (a) If the principal section of the polarizer makes an angle of 30° with one of the principal vibration directions of the plate, what is the ratio of the axes of the elliptically polarized light? (b) What is the inclination of the ellipse relative to the principal section of the polarizer?

9. A Wollaston prism is made of calcite, each oppositely directed half prism having a refracting angle of 30° . What is the angle between the two transmitted beams if the incidence is normal? In a **Wollaston prism** the optic axes in the two half prisms are respectively parallel and perpendicular to their *bases* and refracting edges.

10. Clear Fresnel's normal velocity equation of fractions and solve the resulting biquadratic equation for $2V^2$, verifying the equation on page 208.

11. Express $2V_1^2$ and $2V_2^2$ in terms of the angles g_1 and g_2 , verifying the equations on page 209.

12. Show that the two waves associated respectively with the velocities V_1 and V_2 (problem 10) are vibrating at right angles to each other by proving that $L_1L_2 + M_1M_2 + N_1N_2 = 0$.

Chapter 13. Optical Activity

13.1 Optically Active Materials

Certain materials, such as sugar solutions and quartz along its optic axis, have the property of rotating the plane of polarization. This property is called **optical activity**. The angle through which the plane of polarization is rotated is proportional to the length of path in the active substance and, in solutions, is proportional to the concentration. In the latter case, the rotation per unit length per unit concentration is called the **specific rotatory power**, being usually given in degrees per decimeter and gram per cubic centimeter of solution. Table 13.1 lists the values for various kinds of sugar for sodium light at a temperature of 20°C, the symbol c being the concentration in grams per cubic centimeter of solution. Optical activity provides a means for the quantitative analysis of solutions of active substances.

TABLE 13.1

Sugars	Specific Rotatory Power
Sucrose (cane)	$+66.462 + 0.870 c - 2.35 c^2$
Dextrose	$+52.50 + 1.88 c$
Fructose	$-113.96 + 25.8 c$
Maltose	$+138.48 - 1.837 c$
Lactose	$+52.53$
Invert sugar	$-19.447 - 6.068 c$

Specific rotatory power depends on wave length in a manner similar to that of index of refraction. Hence a formula of the Cauchy type is usually fairly accurate:

$$[\alpha] = A + \frac{B}{\lambda^2}$$

$[\alpha]$ being the specific rotatory power. Furthermore, it depends not only on the temperature, but also on the concentration, as shown in the table. In other words, the relation between rotation and concentration is almost, but not quite, linear. This effect is

sometimes ignored, or one attempts to eliminate it by using nearly standard concentrations, such as 0.26 gm/cm^3 ,³ in the case of sucrose. The temperature and wave length are generally specified by a superscript and a subscript attached to the bracket around the specific rotatory power. Thus one may write for the rotation α due to a length l of solution of concentration c :

$$\alpha = [\alpha]_{5893}^{20^\circ\text{C}} lc$$

Fresnel showed that optical activity is accompanied by a circular double refraction, meaning that in such materials right- and left-hand circularly polarized light travels with different velocities. He constructed a composite prism as shown in Fig. 13.1, using the



FIG. 13.1. Fresnel prism for demonstrating circular double refraction along the optic axis of quartz.

two varieties of quartz which rotate equally in opposite directions. With it he was able to separate an incident monochromatic plane polarized beam into two which were circularly polarized in the opposite sense. It is easy to repeat this experiment with a Fresnel prism, a spectrometer, quarter-wave plate, and Nicol prism. Later on, von Fleishl performed a similar experiment, using prismatic cells to hold solutions of sugar which rotate in opposite senses.

The splitting of the beam into circularly polarized components is usually far too small to observe directly, but the phase difference may be the cause of an appreciable rotation of the plane of polarization. In Fig. 13.2(a), the linear incident vibration is resolved into two oppositely rotating vectors which are in phase at A and represent the circularly polarized components. On emerging, the relative phase is as shown in Fig. 13.2(b), where the angle

$$AOB = \frac{2\pi}{\lambda} (n_r - n_l)l$$

the resultant becoming a linear vibration rotated through the angle $\alpha = \frac{1}{2}AOB$.

The difference in velocity of the two circularly polarized components can be shown to be a result of asymmetry in structure.

In active materials the crystal cell or molecule can have no center of symmetry or plane of symmetry. The proof will be omitted here. In some materials, the molecules themselves are too symmetrical to cause optical activity in a liquid or amorphous form, but the required asymmetry arises from the structure of the crystal. For example, crystalline quartz is active, but quartz glass is not. The crystal structure of quartz shows a spiralling of the

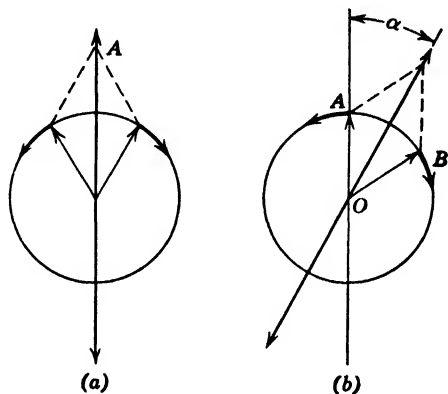


FIG. 13.2. Rotation of the plane of polarization. (a) The entering plane polarized vibration may be resolved into oppositely rotating circularly polarized vibrations. (b) The emerging components combine to give a resultant which is rotated through an angle α , the phase difference being 2α .

molecules of SiO_2 , sometimes in the sense of a right-hand screw and sometimes in the opposite sense. One of these is right hand or dextrorotatory, whereas the other is left hand or levorotatory. There is some ambiguity in these terms. In the case of optical activity of crystals, it is customary to designate a rotation as right-handed if the turning is related to the direction of propagation of light in the same sense as the rotation and translation of a right-hand screw. On the other hand, in solutions the viewpoint is usually that of the observer toward whom the light is travelling, and by a right-hand rotation is meant a clockwise one.

Optically active solutions or liquids usually contain molecules in which there are atoms of carbon with their four valences satisfied by dissimilar atoms or groups of atoms. The simplest example is lactic acid, whose molecule may be thought to consist of a tetra-

hedral arrangement of its parts as shown in Fig. 13.3. There are two possible forms, which are like mirror images and rotate in opposite senses. They are called enantiomorphic isomers. Often one form is produced to the exclusion of the other. If a molecule contains two active carbon groups, there may be three arrangements called **optical isomers**. If both carbons cause a right-hand rotation, the compound has the *dextro* form; if both carbons cause a left-hand rotation, it is *levo*; while if they compensate one another, the material is said to have a *meso* form. The active molecules contain neither a center of symmetry nor a plane of symmetry.

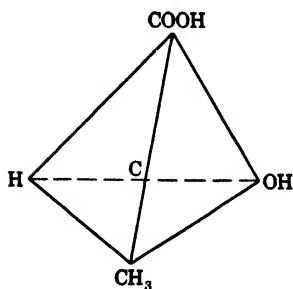


FIG. 13.3. Molecular structure of lactic acid, showing its asymmetry.

13.2 Polarimetry

In order to measure angles of rotation, the analyzer of the polariscope must be mounted on a graduated circle. The instrument is then called a **polarimeter**. Since it is difficult to set accurately on the extinction point, a so-called “**half shade**” is inserted in front of the polarizer. The half shade causes light to vibrate in slightly different directions in the two halves of the sharply divided field of view. The extinction points for the two halves are usually 1° to 2° apart, and one sets between them for an equal intensity. This can be done quite accurately (Experiment 5).

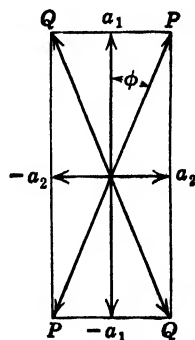
One kind of half shade consists of a small Nicol prism covering half the polarizing Nicol. Each Nicol transmits vibrations in its principal plane, and these are set at the required small angle with each other. This type of half shade is the Lippich. Another is the Cornu-Jellet prism, in which the polarizer is cut lengthwise, a narrow wedge-shaped portion removed, and the parts cemented together. In another type, the biquartz, two plates of quartz of the *dextro* and *levo* form are mounted in juxtaposition in front of the polarizer. They are cut perpendicularly to the optic axis and ground to such a thickness that the plane of polarization is rotated almost 90° in opposite directions in the two halves. For the principal Fraunhofer lines, quartz gives the rotations per millimeter shown in Table 13.2. Accordingly, the biquartz gives its

optimum accuracy only with the wave length for which it was designed.

TABLE 13.2

Fraunhofer line	A	B	C	E	F	G	K
Degrees per millimeter	12.67	17.32	21.70	27.54	32.77	42.60	52.15

The Laurent plate, another type of half shade, consists of a half-wave plate of quartz or mica. It is set so that one of its principal vibration directions a_1 , Fig. 13.4, makes a small angle ϕ with the polarizer principal plane PP . The light it transmits is plane polarized, QQ , and makes twice this small angle with the original vibration, Art. 12.11. The plate covers half the field, the other half often being filled in with a glass plate of about the same retardation and dispersive power to obtain a disappearing boundary when the fields are matched.



PROBLEMS

1. A column of sucrose solution 20 cm long is found to rotate the plane of polarization of sodium light through an angle of 10° . What is the concentration of the solution in grams per liter?

2. The specific rotatory power of sucrose for sodium light is 66.5 degrees per decimeter and gram per cubic centimeter of solution. What is the magnitude of the circular birefringence of a solution containing 26 gm of sucrose per 100 cm^3 of solution?

3. (a) What thickness of quartz cut perpendicularly to its axis will rotate plane polarized sodium light through 90° ? (b) What will be the difference in angle of rotation of this same plate for the F and C Fraunhofer lines?

4. A solution is made up by dissolving a 26.00-gm sample of sucrose in enough water to give 100 cm^3 of solution. The plane of polarization of sodium light is rotated through 34.15° by this solution in a tube 200 mm long. What is the purity of the sucrose if none of the impurities is optically active?

5. What thickness of quartz should be used in making a Laurent half shade for sodium light, the thickness to be not less than 1 mm? To provide a disappearing boundary at balance, one should fit a glass plate of nearly the same dispersive power and retardation next to the quartz plate. What kind of glass, selected from the tables in the Appendix, should be used, and what should be its thickness? The dispersive power of quartz is 0.0129.

FIG. 13.4. Laurent plate. The horizontal component a_2 of the incident vibration PP is retarded by 180° relatively to the vertical component a_1 in passing through the half-wave plate. The resultant vibration is then in the direction QQ .

Chapter 14. Refraction and Dispersion

14.1 Theory of Refraction

It was pointed out that, according to the electromagnetic theory, the index of refraction n is equal to the square root of the dielectric constant ϵ , and that both these quantities vary with the frequency of the radiation. According to the definition of the dielectric displacement D ,

$$D \equiv E + 4\pi P$$

and

$$D = \epsilon E$$

where P , the polarization, is the electrical dipole moment per unit volume, and E is the electric field strength. It is evident that

$$n^2 = \epsilon = 1 + 4\pi \left(\frac{P}{E} \right)$$

Thus a theory of refraction or **dispersion**, which is the variation of refraction with frequency, reduces to a theory of the nature and cause of dielectric polarization P . For molecules having no intrinsic dipole moment, this is the vector sum of the product of the charges and their displacements from equilibrium.

The polarization is proportional to the average field F existing in the medium. This in itself depends on the polarization approximately as

$$F = E + \alpha P$$

where α may be proved to be $4\pi/3$ for a cubic crystal and for a completely fortuitous distribution of dipoles. At first, assuming only one kind of molecule with a polarization $p = \beta F$ for each molecule, the moment per unit volume becomes

$$P = N\beta(E + \alpha P)$$

so that

$$\frac{P}{E} = \frac{\epsilon - 1}{4\pi} = \frac{N\beta}{1 - N\beta\alpha}$$

where N is the number of molecules per unit volume. Thus one finds that

$$4\pi N\beta = (\epsilon - 1) - N\beta\alpha(\epsilon - 1)$$

or

$$N\beta[4\pi + \alpha(\epsilon - 1)] = \epsilon - 1$$

The number of molecules per gram is

$$\frac{N}{\rho} = \frac{N_0}{M}$$

where ρ is the density, N_0 is Avogadro's number, and M is the molecular weight. Then

$$\frac{4\pi N_0 \rho \beta}{M} = \frac{\epsilon - 1}{1 + \frac{\alpha}{4\pi}(\epsilon - 1)}$$

Thus, for a given molecular compound in any state: liquid, gas, or solid, one finds that

$$\frac{1}{\rho} \left(\frac{\epsilon - 1}{1 + \frac{\alpha}{4\pi}(\epsilon - 1)} \right) = \frac{4\pi N_0 \beta}{M}, \text{ a constant}$$

If $\alpha = 4\pi/3$, this reduces to the simpler form

$$\frac{1}{\rho} \left(\frac{\epsilon - 1}{\epsilon + 2} \right) = \text{constant}$$

or

$$\frac{1}{\rho} \left(\frac{n^2 - 1}{n^2 + 2} \right) = \text{constant}$$

the constant being called the **specific refractivity**. A knowledge of specific refractivity enables one to find ϵ or n for a given substance in various states of different density. Debye extended this theory to include the case of molecules which have a dipole moment even in zero field. The refractivity is then found to consist of a constant term plus a term which is proportional to the reciprocal of the absolute temperature. Thus

$$\left(\frac{\epsilon - 1}{\epsilon + 2} \right) \frac{1}{\rho} = \frac{a}{T} + b$$

where $a = 4\pi N_0 \mu^2 / 9Mk$ enables one to determine the dipole moment μ of a molecule from the temperature variation of its refractivity, k being Boltzmann's constant.

14.2 Refractivity

The refractivity R of a mixture of s substances or molecules may be readily found from the refractivities r_1, r_2, \dots, r_s of the components whose mass per gram of the mixture may be represented by m_1, m_2, \dots, m_s . Then

$$\frac{n^2 - 1}{n^2 + 2} \cdot \frac{1}{\rho} = R = m_1 r_1 + m_2 r_2 + \dots + m_s r_s$$

which is readily justified by noting that one must sum up terms like

$$r_1 \rho_1 = \frac{4\pi N_0 \beta_1}{3M_1} \rho_1$$

over all the molecules to obtain the total

$$\frac{n^2 - 1}{n^2 + 2} = r_1 \rho_1 + r_2 \rho_2 + r_3 \rho_3 + \dots + r_s \rho_s \equiv R\rho$$

Even in the case of a compound, the refractivity as computed from the refractivities of the elements using the mixture formula sometimes gives satisfactory results. If the atomic weights are a_1, a_2, a_3, \dots and the number of atoms per molecule are $\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3, \dots$, the molecular weight is $M = \mathbf{n}_1 a_1 + \mathbf{n}_2 a_2 + \mathbf{n}_3 a_3 + \dots$ and the mass of each element per unit mass of compound is

$$m_1 = \frac{\mathbf{n}_1 a_1}{M}; \quad m_2 = \frac{\mathbf{n}_2 a_2}{M}; \quad \text{etc.}$$

Thus by the mixture formula

$$RM = \mathbf{n}_1(a_1 r_1) + \mathbf{n}_2(a_2 r_2) + \dots$$

where the $a \cdot r$ terms are called the atomic refractivities, and RM is the molecular refractivity.

It is not always possible to obtain good results by considering a chemical compound to be a mixture of elements. The computation may be refined, without becoming too complicated, by using a different refractivity according to the valence of the ion, and by adding corrections for the number of bonds and for "conjugation"

of bonds. When this is done, the results are generally satisfactory except in the case of polar molecules, for which one obtains only the temperature-independent portion of the refractivity.

14.3 Dispersion

To investigate dispersion it is required to determine how the polarization p of each atom in the expression

$$n^2 = \epsilon = 1 + 4\pi \frac{P}{E} = 1 + 4\pi \sum \frac{p}{E}$$

varies with the frequency of the incident light. Thus it is necessary to develop p/E as a function of the frequency ν . This problem involves the interaction of a light wave with the electrons in an atom. It is well known that such problems are outside the scope of classical theory and must be solved by quantum mechanics. However, one can find the functional form of the above expression most quickly by classical theory, even though the constants must be altered in regard to magnitude and physical interpretation, as will be shown.

In the electron theory of Lorentz and Drude one writes

$$P = \sum_h N_h e_h \mathbf{r}_h$$

where e_h , N_h , and \mathbf{r}_h are respectively the charge, number per cubic centimeter, and displacement from equilibrium of the h th electron or ion. We may find \mathbf{r} from the differential equation for forced vibrations of these particles:

$$m_h \frac{d^2 \mathbf{r}_h}{dt^2} + h \frac{d\mathbf{r}_h}{dt} + k_h \mathbf{r}_h = e_h E$$

where k is a force constant and h is a damping constant. The value of E in the incident wave may be written as

$$E = E_0 e^{j\omega t}$$

where $\omega = 2\pi\nu$. By substituting the solution $\mathbf{r} = \mathbf{r}_0 e^{j\omega t}$, one finds the relation

$$(-m_h \omega^2 + jh_h \omega + k_h) \mathbf{r}_h = e_h E$$

Thus

$$\frac{P}{E} = \sum_h \frac{N_h e_h^2}{k_h + jh_h \omega - m_h \omega^2}$$

which may be put into the following form

$$\frac{P}{E} = \sum_h \frac{N_h e_h^2}{m_h(\omega_{0h}^2 + j\omega_h' \omega - \omega^2)}$$

where

$$\omega_{0h} \equiv \sqrt{\frac{k_h}{m_h}}$$

is the resonance angular frequency for the h th particle. The significance of the complex term is that there is an absorption. This can be proved as follows. Define a quantity κ by the equation

$$\mathbf{n} = \sqrt{\epsilon} = n(1 - j\kappa)$$

Then for an electromagnetic wave

$$\frac{1}{V} = \frac{\sqrt{\epsilon}}{c} = \frac{n - jn\kappa}{c}$$

so that

$$E = E_0 e^{j\frac{2\pi}{T}\left(t - \frac{z}{V}\right)} = E_0 e^{j\frac{2\pi}{T}\left(t - \sqrt{\epsilon} \frac{z}{c}\right)}$$

or

$$E = E_0 e^{j\frac{2\pi}{T}\left[t - \frac{z}{c}(n - jn\kappa)\right]} = E_0 e^{-\frac{2\pi}{\lambda} n\kappa z} e^{j2\pi\left(\frac{t}{T} - \frac{nz}{\lambda}\right)}$$

This shows that there is an exponential reduction in amplitude as the wave progresses in the z direction, the absorption coefficient for intensity I in the equation $I = I_0 e^{-az}$ being

$$a = \frac{4\pi n\kappa}{\lambda}$$

The complex dielectric constant is given by the expression

$$\epsilon = n^2(1 - j\kappa)^2 = 1 + 4\pi \sum_h \frac{N_h \frac{e_h^2}{m_h}}{\omega_{0h}^2 + j\omega_h' \omega - \omega^2}$$

On separating this into real and imaginary parts, one finds that

$$n^2(1 - \kappa^2) = 1 + 4\pi \sum_h \frac{N_h \frac{e_h^2}{m_h} (\omega_{0h}^2 - \omega^2)}{(\omega_{0h}^2 - \omega^2)^2 + \omega_h'^2 \omega^2}$$

and

$$2n^2\kappa = 4\pi \sum_h \frac{N_h \frac{e_h^2}{m_h} \omega_h' \omega}{(\omega_{0h}^2 - \omega^2)^2 + \omega_h'^2 \omega^2}$$

Graphs of the optical constants n and κ computed from these equations have the form shown in Fig. 14.1. The constant κ is usually

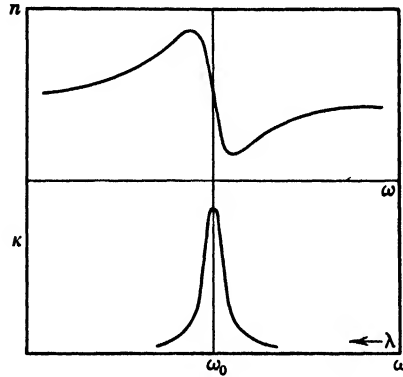


FIG. 14.1. Variation of index of refraction n and index of absorption κ with angular frequency ω in the neighborhood of an absorption band.

called the **index of absorption**, while n is the **index of refraction**. When the medium is an insulator, ω' is negligible unless ω is close to an absorption frequency ω_0 . Thus at some distance from the region of absorption one finds the simpler results

$$\kappa \doteq 0$$

$$n^2 \doteq 1 + 4\pi \sum_h \frac{N_h \frac{e_h^2}{m_h}}{\omega_{0h}^2 - \omega^2}$$

A graph showing the variation of n^2 with angular frequency in the vicinity of the first absorption band is shown in Fig. 14.2. Of course, the results become inaccurate as ω approaches ω_1 . On passing through each absorption band, the value of n^2 should decrease by the amount

$$\frac{4\pi N_h e_h^2}{\omega_{0h}^2 m_h}$$

for the particular resonating particle, as shown for one particle in Fig. 14.2. That this is not found to be the case quantitatively indicates a shortcoming of the classical theory. At very high frequencies the equation for n^2 shows that the index of refraction should approach unity from the lower side. The fact that the

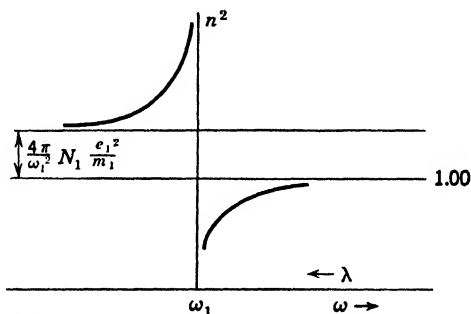


FIG. 14.2. Change in the magnitude of n^2 on passing through an absorption band.

index of refraction of x-rays is less than one is in agreement with this expectation.

14.4 Quantum Theory of Dispersion

Instead of imagining forced vibrations of electrons, one must think of induced transitions between quantized energy states. This relates the theory of dispersion to that of scattering. The development of this theory is beyond the scope of this book, but it is definitely of interest to compare the quantum-theory dispersion formula with that obtained by the classical theory. According to Kramers, the index of refraction n may be expressed by the formula

$$n^2 - 1 = \sum_i \frac{N_i c^3}{8\pi^3} \left\{ \sum_k \frac{a_{i \rightarrow k}}{\nu_{ik}^2 (\nu_{ik}^2 - \nu^2)} - \sum_g \frac{a_{i \rightarrow g}}{\nu_{ig}^2 (\nu_{ig}^2 - \nu^2)} \right\}$$

where the subscripts i , k , and g refer to energy levels, k being higher than i , and g lower than i . The number of atoms per unit volume in the energy state i is denoted by N_i . The symbols $a_{i \rightarrow k}$ and $a_{i \rightarrow g}$ are respectively the probabilities for upward and downward transitions induced by the radiation. Under ordinary conditions, most of the atoms are in the ground state, so that the

second term is negligible. If one considers the neighborhood of one particular absorption frequency ν_{ik} , as for example in the refractive index of sodium vapor near the D lines, one may write

$$n^2 - 1 = \frac{Ne^3}{8\pi^3} \frac{a_{i \rightarrow k}}{\nu_{ik}^2(\nu_{ik}^2 - \nu^2)}$$

This is the quantum theory equivalent of the corresponding classical formula

$$n^2 - 1 = \frac{Ne^2}{\pi m(\nu_h^2 - \nu^2)}$$

A comparison of the two formulas shows that

$$\frac{e^2}{m} \text{ corresponds to } \frac{c^3}{8\pi^2} \frac{a_{i \rightarrow k}}{\nu_{ik}^2}$$

Thus one may artificially "correct" the classical formula by the use of *effective* values of e^2/m instead of actual values. For example, the experiments of Füchtbauer and Hoffman on the dispersion of sodium vapor near the D lines show that the effective value of e^2/m is only 0.14 times the actual value. From this result, one may compute the transition probability a_{1S-2P} . The value agrees well with that computed from the wave functions of the sodium atom and provides a good check of the quantum-theory formula.

14.5 X-ray Dispersion

In the x-ray region one finds absorption *edges* instead of absorption *lines*, Art. 18.1. One may divide each continuous absorption band, which has a low frequency limit, into a series of strips and consider the electrons associated with each edge to be distributed among the resulting elementary strips. By integrating the electron contributions to p/E , Kramers, Kallman, and Mark first derived an expression which describes x-ray dispersion quite well. However, the assumed λ^3 distribution law and the assumed effective number of electrons in the K and L levels are not always valid. An improved formula which takes these matters into consideration was derived by Hönl¹ for the neighborhood of the K edge. This is

¹ HÖNL, *Zeitschrift für Physik*, **84**, 1-16 (1933).

$$\delta_k \equiv 1 - n = \frac{N_a e^2}{2\pi m \nu^2} \frac{2^7 e^{-4}}{9} \left\{ \frac{4}{(1 - \Delta_k)^2} (1 + x^2 \ln |1 - x^{-2}|) - \frac{1}{(1 - \Delta_k)^3} \left(\frac{2}{3} + 2x^2 + x^3 \ln \left| \frac{1 - x}{1 + x} \right| \right) \right\}$$

where $x = \nu_k/\nu$, and N_a is the number of atoms per unit volume. The quantity Δ_k is given by

$$\Delta_k = \frac{(Z - s)^2 + \frac{1}{4} \alpha^2 (Z - s)^4 - \frac{h\nu_k}{Rhc}}{(Z - s)^2}$$

where Z is the atomic number, s is the screening constant for the K level, α is Sommerfeld's fine-structure constant, R is Rydberg's constant, h is Planck's constant, and c is the speed of light. Figure

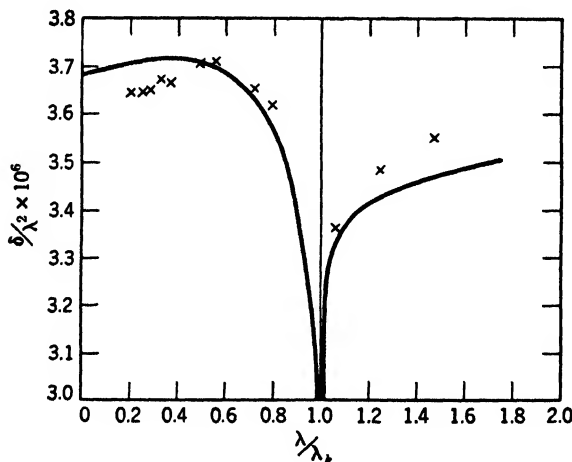


FIG. 14.3. X-ray dispersion in the neighborhood of an absorption edge at $\lambda = \lambda_k$. The curve is from the wave-mechanical theory by Hönl. The crosses are experimental values for quartz as measured by Larsson.

14.3 shows a comparison of this formula with the experimental data of Larsson, the curve being computed by Hönl's formula and the crosses being experimental values.

PROBLEMS

1. (a) Compute the refractivities of N_2 and of CO_2 at normal temperature and pressure. (b) What is the index of refraction of a mixture of the two gases containing 20 per cent by weight of CO_2 ?

2. Compute the index of refraction of N_2O at normal temperature and pressure, considering it to be a mixture of the two gases instead of a chemical compound. What is the discrepancy between the computed and measured values?

3. The values of $(n - 1)10^7$ for hydrogen at normal temperature and pressure are 1759.96 at a wave length of 1854.64 Å, and 1396.50 at 5462.26 Å. Assuming normal dispersion and a single principal absorption line in the ultra-violet, determine the effective value of e^2/m and compare it with the actual value.

4. (a) How does the refraction of x-rays change the form of Bragg's formula? (b) If the apparent lattice constant of calcite is 3029.04 XU (1 XU = 10^{-11} cm) for first order diffraction, while for the third order it is 3029.40 XU, what is the index of refraction of calcite for x-rays having a wave length of 1000 XU?

5. (a) What is the critical angle for total reflection of x-rays having a wave length of 708 XU by a glass surface for which $(n - 1) = 1.64 \times 10^{-6}$? (b) What will be the maximum deviation of such a totally reflected beam at a distance of 25 cm?

Chapter 15. Optical Constants of Metallic Absorbers

15.1 Metallic Media

The absorption of the longer wave lengths in metals is due to the Joule effect associated with conduction electrons. The contribution to absorption due to photoelectric excitation and ionization takes place at short wave lengths and will not be considered, except in so far as it affects dielectric polarization.

The absorption in metals is generally so great that the amplitude drops to a small fraction of its value in a distance of less than a wave length. One should accordingly distinguish between planes of equal phase and of equal amplitude, as they may be quite different. This discussion is restricted to plane waves. The plane of constant amplitude is practically always parallel to the surface regardless of the angles of incidence and refraction. It will be found that this fact leads to a variation in the index of refraction or wave velocity, and of absorption coefficient, with the angle of incidence. The values usually tabulated are for perpendicular incidence and cannot be applied to oblique incidence without further investigation.

For metallic media, the fundamental equations, Art. 12.2, become

$$\text{curl } E = - \frac{1}{c} \frac{\partial H}{\partial t} \quad (1)$$

$$\text{curl } H = \frac{1}{c} \left(4\pi\sigma E + \epsilon \frac{\partial E}{\partial t} \right) \quad (2)$$

$$\text{div } E = 0 \quad (3)$$

$$\text{div } H = 0 \quad (4)$$

where σ is the conductivity of the metal, and the other symbols have their usual significance. The magnetic permeability is again taken equal to unity at optical frequencies. The significance of the vector notation is made clear by comparing these equations with those given in Art. 12.2.

Differentiating (2) with respect to the time t and eliminating H by the use of (1) lead one to the wave equation

$$\frac{\epsilon}{c^2} \frac{\partial^2 E}{\partial t^2} + \frac{4\pi\sigma}{c^2} \frac{\partial E}{\partial t} = \frac{\partial^2 E}{\partial x^2} + \frac{\partial^2 E}{\partial y^2} + \frac{\partial^2 E}{\partial z^2}$$

Choose the xy plane as the surface of the metal, z being then the normal, and let the plane of incidence be the xz plane. Then for plane waves $\partial/\partial y = 0$.

15.2 Optical Constants

A particular solution of the wave equation under these conditions is

$$E = E_0 e^{j\omega[t - (ax + bz)]}$$

where

$$\omega = 2\pi\nu = \frac{2\pi}{T}$$

Substituting this into the wave equation to determine the necessary relation between the constants, one finds that

$$\frac{1}{c^2} (-\epsilon\omega^2 + j4\pi\sigma\omega) = -\omega^2(a^2 + b^2)$$

so that

$$a^2 + b^2 = \frac{1}{c^2} (\epsilon - j2\sigma T)$$

When the conductivity σ is not zero, as in the case of metals, a and b cannot both be real. For the general case, they may be written in the complex form

$$\begin{aligned} a &= \frac{\sin r}{V} - j \frac{\kappa_1}{V} \quad \text{or} \quad \frac{n \sin r}{c} - j \frac{n\kappa_1}{c} \\ b &= \frac{\cos r}{V} - j \frac{\kappa_2}{V} \quad \text{or} \quad \frac{n \cos r}{c} - j \frac{n\kappa_2}{c} \end{aligned}$$

The equation of the wave is then

$$E = E_0 e^{-\frac{2\pi n}{\lambda} (\kappa_1 x + \kappa_2 z)} e^{j\omega \left\{ t - \frac{n(x \sin r + z \cos r)}{c} \right\}}$$

which is seen to be the equation of a plane wave, since the phase is constant over the plane

$$x \sin r + z \cos r = \text{constant}$$

However, the amplitude varies with x and z , being constant over a different plane

$$\kappa_1 x + \kappa_2 z = \text{constant}'$$

which may be written in the form

$$\kappa(x \sin \alpha + z \cos \alpha) = \text{constant}'$$

where κ is the index of absorption. In these equations r is the angle with the z axis of the normal to the plane of constant phase, while α is a similar angle made by the normal to the plane of constant amplitude. In materials of high absorption, the locus of points of equal amplitude is practically parallel to the surface, so that in this case α may be put equal to zero. The wave, writing only the real part, is given by

$$E = E_0 e^{-\frac{2\pi n \kappa}{\lambda} z} \cos \frac{2\pi n}{\lambda} \{Vt - (x \sin r + z \cos r)\}$$

The relation between n , κ , α , r , ϵ , and σ is obtained from

$$a = \frac{1}{c} (n \sin r - j n \kappa \sin \alpha)$$

$$b = \frac{1}{c} (n \cos r - j n \kappa \cos \alpha)$$

since the wave equation requires that

$$a^2 + b^2 = \frac{1}{c^2} \{\epsilon - j2\sigma T\}$$

It follows that

$$\epsilon' \equiv \epsilon - j2\sigma T = n^2(1 - \kappa^2) - j2n^2\kappa \cos(r - \alpha)$$

By equating the real parts and the imaginary parts, one obtains

$$n^2(1 - \kappa^2) = \epsilon$$

and

$$2n^2\kappa \cos(r - \alpha) = 2\sigma T$$

The second equation shows that n and κ depend on $r - \alpha$, which is practically equal to r since α is very small. For perpendicular incidence, $r = 0$, the relations are

$$\begin{aligned}n_0^2(1 - \kappa_0^2) &= \epsilon \\2n_0^2\kappa_0 &= 2\sigma T\end{aligned}$$

which are usually given without calling attention to the fact that they apply only to normal incidence.

In the case of normal incidence the wave becomes

$$E = E_0 e^{-\frac{2\pi n_0 \kappa_0}{\lambda} z} \cos 2\pi \frac{n_0}{\lambda} (V_0 t - z)$$

so that the intensity decreases exponentially as

$$I = I_0 e^{-\frac{4\pi n_0 \kappa_0}{\lambda} z}$$

The absorption coefficient $4\pi n_0 \kappa_0 / \lambda$ is very large for most metals.

In the more general case of $r \neq 0$ one finds that

$$E = E_0 e^{-\frac{2\pi n \kappa}{\lambda} z} \cos \omega \left[t - n \frac{x \sin r + z \cos r}{c} \right]$$

and consequently

$$I = I_0 e^{-\frac{4\pi n \kappa}{\lambda} z}$$

The relations between the constants for normal and for oblique incidence are

$$\begin{aligned}n^2(1 - \kappa^2) &= n_0^2(1 - \kappa_0^2) \\n^2 \kappa \cos r &= n_0^2 \kappa_0\end{aligned}$$

To calculate n and κ from these and to find how they vary with the angle of incidence, one introduces Snell's law, according to which

$$n^2 \sin^2 r = \sin^2 i$$

and

$$n^2 \cos^2 r = n^2 - \sin^2 i$$

Making use of the latter, it is readily found that

$$n^4 \kappa^2 - n^2 \kappa^2 \sin^2 i = n_0^4 \kappa_0^2$$

and, of course,

$$n^2(1 - \kappa^2) = n_0^2(1 - \kappa_0^2)$$

These two equations may be solved simultaneously for $2n^2$ and $2n^2\kappa^2$, giving

$$2n^2 = \{[n_0^2(1 - \kappa_0^2) - \sin^2 i]^2 + 4n_0^4\kappa_0^2\}^{1/2} \\ + \{n_0^2(1 - \kappa_0^2) + \sin^2 i\}$$

and

$$2n^2\kappa^2 = \{[n_0^2(1 - \kappa_0^2) - \sin^2 i]^2 + 4n_0^4\kappa_0^2\}^{1/2} \\ - \{n_0^2(1 - \kappa_0^2) - \sin^2 i\}$$

Conversely,

$$2n_0^2 = \{n^4(1 + \kappa^2)^2 - 4n^2\kappa^2\sin^2 i\}^{1/2} + n^2(1 - \kappa^2)$$

and

$$2n_0^2\kappa_0^2 = \{n^4(1 + \kappa^2)^2 - 4n^2\kappa^2\sin^2 i\}^{1/2} - n^2(1 - \kappa^2)$$

These may be reduced to Ketteler's formulas by introducing the quantities

$$m \equiv n \cos r \quad \text{and} \quad K^2 \equiv n^2\kappa^2$$

giving

$$m^2 - K^2 + \sin^2 i = n_0^2 - K_0^2$$

and

$$mK = n_0K_0$$

where $K_0 \equiv n_0\kappa_0$. Using these, one finds that

$$2n_0^2 = \{(m^2 - K^2 + \sin^2 i)^2 + 4m^2K^2\}^{1/2} + (m^2 - K^2 + \sin^2 i)$$

and

$$2K_0^2 = \{(m^2 - K^2 + \sin^2 i)^2 + 4m^2K^2\}^{1/2} - (m^2 - K^2 + \sin^2 i)$$

To make clear the order of magnitude of the quantities in a typical case and to show how they vary with i , the following results by Wilsey¹ for gold are quoted. That the amplitude de-

Gold, $\lambda = 0.620 \mu$

i	κ_i	n_i
0°	10.03	0.3104
20°	10.15	0.469
40°	10.20	0.722
60°	10.85	0.916
70°	10.86	0.985
80°	11.04	1.033
85°	11.72	1.032

¹ WILSEY, *Physical Review*, **8**, 391 (1916).

creases very rapidly with z can be seen by the equation for gold at normal incidence

$$E = E_0 e^{-18z, \lambda}$$

In only one eighteenth of the wave length, the amplitude is $1/e$ of the incident value, so that after penetrating to a distance of only one wave length one finds that

$$E = \frac{E_0}{e^{18}} = \frac{E_0}{65,000,000}$$

Moreover, the intensity varies as E^2 . In general, metals are the best absorbers of light that are known. One can appreciate that a direct measurement of the optical constants of metals would be extremely difficult. Yet this has been done by Kundt, although the results are not very accurate because of the difficulty in obtaining uniform films and prisms of small angle, and in the measurement of their thickness and angle. It is interesting that one frequently finds that the wave velocity, $V = c/n$, is greater than c , since n is often less than unity. However, this is a phase velocity, as one can see by the derivation of $n = c/V$ by Huygens' principle. One always finds that the group or signal velocity is less than c , as is to be expected. Tables of optical constants of metals are given by Valasek in the *International Critical Tables*, 5, 248-256.

In testing the theoretical relation between optical constants and conductivity

$$n_0^2 \kappa_0 = \sigma T$$

one finds that the agreement is good only at very long wave lengths. The value of the conductivity which should be used is that which applies to the same frequency for which n_0 and κ_0 are measured. When this is done, or even approached in many cases, the agreement is good. In many instances one does not have to go very far into the infrared with measurements of the optical constants to find that the ordinary direct-current conductivity may be used for σ in the preceding formulas.

15.3 Normal Reflection

The Fresnel formulas for reflection, Art. 12.3, hold for metals if one replaces the index of refraction by its complex value $n_0 - j n_0 \kappa_0$. This follows from the fact that the fundamental equations

have the same mathematical form in ϵ' as those for an insulator, and for normal incidence

$$\epsilon' = \epsilon - j2\sigma T = n_0^2 - j2n_0^2\kappa_0 - n_0^2\kappa_0^2 = (n_0 - jn_0\kappa_0)^2$$

The reflected wave is then given by

$$R = R e^{j\omega\left(t + \frac{z}{c}\right)}$$

where

$$R = -\frac{n-1}{n+1}E$$

as in the insulator case, except that n is now a complex quantity. Thus for metals

$$R = \frac{1 - n_0 + jn_0\kappa_0}{1 + n_0 - jn_0\kappa_0}E = \frac{1 - n_0^2 - n_0^2\kappa_0^2 + j2n_0\kappa_0}{(1 + n_0)^2 + n_0^2\kappa_0^2}E$$

Any complex quantity such as R may be written in the form

$$R = A + jB = \sqrt{A^2 + B^2} e^{j\phi}$$

where

$$\phi = \tan^{-1} \frac{B}{A}$$

On carrying out this operation, it is found that the phase shift on reflection ϕ may be computed by

$$\tan \phi = \frac{2n_0\kappa_0}{1 - n_0^2 - n_0^2\kappa_0^2}$$

The phase difference between the incident wave and the reflected wave is decreased by an amount less than π at any distance $-z$ above the reflecting surface. This displaces the nodes of the standing waves toward the surface so that the first node is at a point below the surface of the metal.

The normal reflectivity \mathcal{R} of the metal surface may be found most readily by multiplying R/E by its complex conjugate, which gives the ratio of the squares of the absolute values of the amplitudes. Thus one finds that

$$\mathcal{R} = \frac{(1 - n_0)^2 + n_0^2\kappa_0^2}{(1 + n_0)^2 + n_0^2\kappa_0^2}$$

Metals usually have a very high reflectivity which approaches unity as the wave length increases. One may write \mathcal{R} in the form

$$\mathcal{R} = 1 - \frac{4n_0}{(1 + n_0)^2 + n_0^2\kappa_0^2}$$

Then, since $n_0^2\kappa_0^2$ is usually greater than one, and n_0 is of the order of unity or less, at least in the *visible* region, one can see that the reflectivity is less than unity by a quantity of the order of magnitude of $4n_0/(4 + n_0^2\kappa_0^2)$, which is usually small. However, the reflectivity drops greatly in the region of shorter waves. In particular, silver, which reflects nearly 96 per cent of the incident light in the visible region, has a reflectivity of only a few per cent at 3200 Å, which rises again at still shorter wave lengths. Thus silver makes a poor mirror for ultraviolet light, but thin films on quartz may be used as a transmission filter for isolating a narrow band of wave lengths in the vicinity of 3200 Å.

15.4 Reflection and Conductivity

Since $n_0^2\kappa_0 = \sigma T$ while $n_0^2 - n_0^2\kappa_0^2 = \epsilon$, one finds that, as the period T increases, the first quantity increases without limit, whereas the second remains relatively constant. Thus for low frequencies or large values of T

$$n_0^2 \rightarrow n_0^2\kappa_0^2 \quad \text{and} \quad n_0 \rightarrow n_0\kappa_0 \rightarrow \sqrt{\sigma T}$$

Then, since $n_0^2(1 + \kappa_0^2)$ becomes large compared with $1 + 2n_0$, one may obtain the limiting expression for \mathcal{R} in the form

$$\mathcal{R} \doteq 1 - \frac{4n_0}{n_0^2(1 + \kappa_0^2)} \rightarrow 1 - \frac{2}{\sqrt{\sigma T}}$$

which should hold approximately at long wave lengths. By eliminating the period T in favor of the wave length λ in microns, and changing the conductivity in units of emu/cm³ to the resistivity r in ohms/cm³, one finds that

$$\frac{1 - \mathcal{R}}{\sqrt{r}} \rightarrow \frac{0.365}{\sqrt{\lambda}}$$

In particular, for $\lambda = 12 \mu$, one finds that the quantity on the right has the value 0.105, while experimental observations give for the quantity on the left the following values:

Ag	0.09	Au	0.138	Ni	0.12
Cu	0.12	Pt	0.106	Steel	0.11

At greater wave lengths the agreement is still better.

15.5 Oblique Reflection

The difficulties in direct measurements of the optical constants of metals have already been pointed out. These constants may be obtained more readily by observations on oblique reflection of polarized light, although in this case the results may be inaccurate because of the presence of surface films, which are hard to avoid.

The Fresnel formulas for oblique reflection may be extended to metals. They are, for insulators,

$$R_p = -E_p \frac{\tan(i-r)}{\tan(i+r)}; \quad R_s = -E_s \frac{\sin(i-r)}{\sin(i+r)}$$

so that

$$\frac{R_p}{R_s} = \frac{\cos(i+r)}{\cos(i-r)} = \frac{\cos i \cos r - \sin i \sin r}{\cos i \cos r + \sin i \sin r}$$

when E_p and E_s are equal to each other.

The general or rigorous discussion of the modification of these equations to metals is quite complicated. Hence the following treatment is restricted to a few important cases where n is so large that 1 or $\sin^2 i$ can be neglected in comparison with n^2 .

Inside the metal, the wave may be represented by the equation

$$D = D_0 e^{j\omega[t - (ax+bz)]}$$

or

$$D = D_0 e^{j\omega[t - (n/c)\{x \sin r + z(\cos r - j\kappa)\}]}$$

which is mathematically of the same form as that for insulators, but with $\cos r$ replaced by $\cos r - j\kappa$. The formula for the amplitude ratio should then apply to metals after making the indicated substitution in Fresnel's formula, the \mathbf{R} 's becoming complex so that

$$\frac{\mathbf{R}_p}{\mathbf{R}_s} = \frac{\cos i \cos r - \sin i \sin r - j\kappa \cos i}{\cos i \cos r + \sin i \sin r - j\kappa \cos i}$$

Eliminating r by Snell's law,

$$n \sin r = \sin i$$

and

$$n \cos r = \sqrt{n^2 - \sin^2 i} \doteq n$$

one finds that

$$\frac{R_p}{R_s} = \frac{R_p e^{j\delta_p}}{R_s e^{j\delta_s}} = \frac{n \cos i - \sin^2 i - jn\kappa \cos i}{n \cos i + \sin^2 i - jn\kappa \cos i}$$

The amplitude ratio R_p/R_s is conveniently represented by $\tan \psi$, where ψ may be measured by a Nicol prism after the phase difference $\Delta = \delta_p - \delta_s$ has been compensated. Then

$$e^{j\Delta} \tan \psi = \frac{(n - \sin i \tan i) - jn\kappa}{(n + \sin i \tan i) - jn\kappa}$$

Multiplying this by its complex conjugate gives

$$\begin{aligned} \tan^2 \psi &= \frac{(n - \sin i \tan i)^2 + n^2 \kappa^2}{(n + \sin i \tan i)^2 + n^2 \kappa^2} \\ &= \frac{n^2(1 + \kappa^2) + \sin^2 i \tan^2 i - 2n \sin i \tan i}{n^2(1 + \kappa^2) + \sin^2 i \tan^2 i + 2n \sin i \tan i} \end{aligned}$$

By putting the expression for $e^{j\Delta} \tan \psi$ into the general complex form $X + jY$, one finds that

$$\tan \Delta = \frac{Y}{X} = \frac{+n\kappa(n - \sin i \tan i) - n\kappa(n + \sin i \tan i)}{n^2 - \sin^2 i \tan^2 i + n^2 \kappa^2}$$

so that

$$\tan \Delta = - \frac{2n\kappa \sin i \tan i}{n^2(1 + \kappa^2) - \sin^2 i \tan^2 i}$$

These equations may be reduced to a simpler form by using the variable α , defined by

$$\tan \alpha = \frac{\sin i \tan i}{n(1 + \kappa^2)^{1/2}}$$

for then it can be shown that

$$\begin{aligned} \cos 2\psi &= (1 + \kappa^2)^{-1/2} \sin 2\alpha \\ \tan \Delta &= -\kappa(1 + \kappa^2)^{-1/2} \tan 2\alpha \end{aligned}$$

From these one may readily compute ψ and Δ for any given pair of optical constants n and κ and for any given angle of incidence i . Typical results are shown in the graphs of Fig. 15.1.

One finds that the phase difference Δ is $-\pi/2$, or $\delta_s - \delta_p = +\pi/2$, at some angle of incidence. This angle is called the principal angle of incidence \bar{i} and may be easily located in the laboratory with the aid of a quarter-wave plate and Nicol prism. The

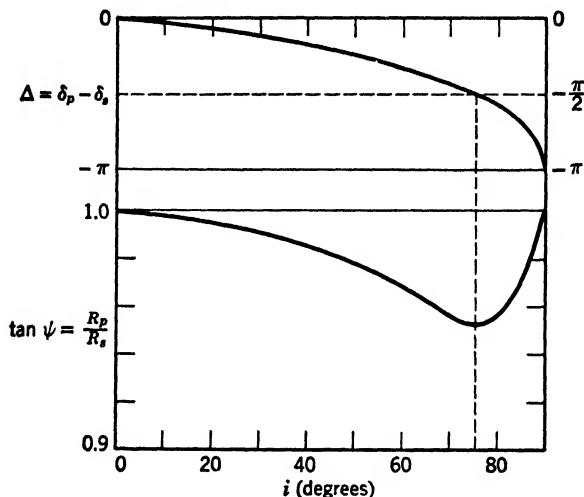


FIG. 15.1. Reflection by metals. Variation with angle of incidence i of the phase difference Δ between reflected p and s components. The lower curve shows the variation with i of the ratio of their amplitudes.

value of ψ at the principal angle of incidence, denoted by $\bar{\psi}$, is called the principal azimuth. It may be found from the extinction setting of the analyzer.

Since $\tan \Delta = \infty$ at the principal angle of incidence, it follows that $\alpha = 45^\circ$ in this case. Then we have the following relations:

$$\sin \bar{i} \tan \bar{i} = n(1 + \kappa^2)^{1/2}$$

$$\cos 2\bar{\psi} = (1 + \kappa^2)^{-1/2}$$

$$\sin 2\bar{\psi} = \kappa(1 + \kappa^2)^{-1/2}$$

From these one obtains the useful formulas

$$\tan 2\bar{\psi} = \kappa$$

and

$$\sin \bar{i} \tan \bar{i} \cos 2\bar{\psi} = n$$

These enable one to measure the optical constants n and κ approximately. The accuracy is limited by the error inherent in locating the angle \bar{i} and by the approximations used in deriving the above formulas.

15.6 Comparison with Insulators

The amplitudes of the reflected components in the case of insulators may also be written in the complex form

$$\mathbf{R}_p = R_p e^{j\delta_p} \quad \text{and} \quad \mathbf{R}_s = R_s e^{j\delta_s}$$

so that

$$\frac{\mathbf{R}_p}{\mathbf{R}_s} = \frac{R_p}{R_s} e^{j\Delta} = e^{j\Delta} \tan \psi$$

Since the amplitude components R_p and R_s for the incident light are considered positive and real, the negative sign in the Fresnel equations for reflection indicates that $\delta_s = \pi$ for all angles of in-

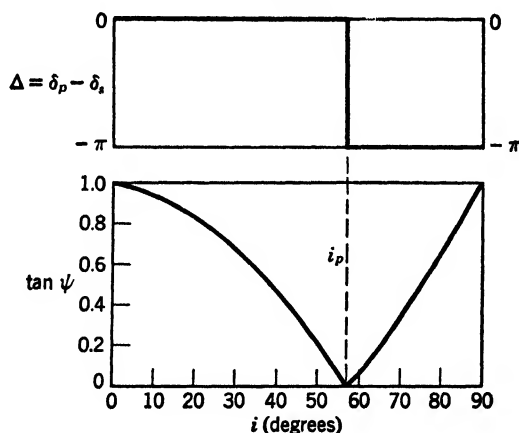


FIG. 15.2. Reflection by an insulator. Same quantities as in Fig. 15.1 to show difference between reflection by metals and by insulators.

cidence, while $\delta_p = \pi$ up to the polarizing angle and is zero beyond. Thus the phase difference $\Delta = \delta_p - \delta_s = 0$ up to the polarizing angle and drops to the value $-\pi$ thereafter, as shown in Fig. 15.2. Simultaneously, the ratio of the amplitude components given by $\tan \psi$ decreases from unity to zero as the angle of incidence approaches the polarizing angle, after which it rises to unity again.

A comparison of these results with those for metals, Fig. 15.1, shows that the discontinuity in Δ is replaced by a gradual decline, whereas the value of $\tan \psi$ never drops to zero but has a rather high minimum. Thus metals do not have a polarizing angle.

In practice, one finds that insulators do not strictly follow the ideal curve obtained by the use of the Fresnel equations, but merely approximate it, the deviations being in the direction of the results obtained for metals. Thus it is found to be extremely difficult, if not impossible, to obtain plane polarized light by reflection from an insulator. Naturally, the best results are obtained when the surface is as perfectly smooth and clean as it is possible to make it. The light reflected at the Brewster angle is then found to be elliptically polarized, but with one axis very much greater than the other. Thus Jamin observed for a freshly split surface of fluorite an axial ratio of 0.0085, while for a surface of clean water this ratio could be reduced to 0.00035.

15.7 Total Reflection

The phenomenon of total reflection inside prisms of glass or other transparent material is often preferred to metallic reflection in many optical instruments. The reflectivity is practically perfect (except for diffraction due to the boundaries) and it does not deteriorate with time as does the reflectivity of metallic mirrors.

Total reflection occurs when the sine of the angle of incidence exceeds the relative index of refraction n . This naturally requires that the relative index of refraction be less than unity, which is the case when the index of the second medium is less than the index of the first. For example, this condition is satisfied when there is internal reflection in glass. Then, if $\sin i$ is greater than n , it follows that

$$\sin r = \frac{\sin i}{n}$$

is greater than unity, and that

$$\cos r = -\frac{j}{n} \sqrt{\sin^2 i - n^2}$$

is purely imaginary, the negative sign of the radical being chosen to obtain an exponential decrease in the amplitude of the wave in the second medium. One might jump to the conclusion that,

since the angle of refraction is complex, there can be no refracted ray and that consequently total reflection takes place. Another method of analysis proves to be more informative. Having been introduced to the use of complex expressions for optical waves, one may substitute the above expressions for $\sin r$ and for $\cos r$ into Fresnel's equations for reflection. The mathematical operations are perfectly legitimate, and the physical significance can best be established after deriving the results. Since the reflected amplitudes will surely be complex, one had best write them in the complex polar form, thus:

$$\frac{R_s}{E_s} e^{j\delta_s} = - \frac{\sin(i - r)}{\sin(i + r)}$$

and

$$\frac{R_p}{E_p} e^{j\delta_p} = - \frac{\tan(i - r)}{\tan(i + r)} = - \frac{\sin(i - r) \cos(i + r)}{\sin(i + r) \cos(i - r)}$$

If one expands these sines and cosines of the sums and differences of two angles and notes that $\cos r$ is imaginary, the following relations are seen to hold between certain functions and their complex conjugates, which are marked with an asterisk. Incidentally, complex conjugates are identical quantities but with the sign of each imaginary term reversed. Thus

$$\sin(i - r) = - \overline{\sin(i + r)^*}$$

$$\cos(i - r) = - \overline{\cos(i + r)^*}$$

and vice versa.

If one multiplies both sides of Fresnel's reflection formulas by their complex conjugates, he finds, because of the above relations, that

$$\left| \frac{R_p}{E_p} \right|^2 = 1 \quad \text{and} \quad \left| \frac{R_s}{E_s} \right|^2 = 1$$

This indicates that the reflectivities are unity and that consequently there is **total reflection**. That a phase shift accompanies total reflection is obvious from the complex form for the amplitudes. Thus total reflection is similar to metallic reflection in that the reflectivity is high and there is a phase shift differing from zero or π .

The difference in phase between the p and s components

$$\Delta = \delta_p - \delta_s$$

is of some interest, since it is utilized in the Fresnel rhomb, a device for producing elliptically or circularly polarized light by two total reflections. From the totality of the reflections it follows that

$$\left| \frac{R_p}{E_p} \right| = 1 \quad \text{and} \quad \left| \frac{R_s}{E_s} \right| = 1$$

so that the ratio of the Fresnel formulas leads to the result:

$$e^{j\Delta} = \frac{\cos(i+r)}{\cos(i-r)}$$

If one writes

$$\cos(i-r) = \rho e^{j\theta}$$

it follows that

$$\cos(i+r) = -\rho e^{-j\theta} = \rho e^{j(\pi-\theta)}$$

Therefore

$$\Delta = \delta_p - \delta_s = \pi - 2\theta$$

so that

$$\tan \frac{\Delta}{2} = \cot \theta$$

To obtain an expression for θ , expand $\cos(i-r)$ and substitute for $\cos r$ and for $\sin r$, obtaining

$$\cos(i-r) = \frac{\sin^2 i}{n} - j \frac{\cos i \sqrt{\sin^2 i - n^2}}{n} = \rho e^{j\theta}$$

Thus one finds that

$$\tan \frac{\Delta}{2} = \cot \theta = - \frac{\sin^2 i}{\cos i \sqrt{\sin^2 i - n^2}}$$

by which one may compute the phase difference for a given angle of incidence in excess of the critical angle. At the critical angle, $\sin i_c = n$, one finds that

$$\tan \frac{\Delta}{2} = \infty, \quad \text{so that} \quad \Delta = \pi$$

Likewise, at grazing incidence, $i = \pi/2$, one also finds that $\Delta = \pi$.² Between these values of i there is a maximum value of Δ , which can be found by differentiating the expression for $\tan(\Delta/2)$ with respect to i . The angle at which this maximum occurs is accordingly given by

$$\sin^2 i_m = \frac{2n^2}{1 + n^2}$$

and at this angle

$$\tan \frac{\Delta_m}{2} = - \frac{2n}{1 - n^2}$$

For air with respect to crown glass, $n = 1/1.52$, one finds, for example, that $i_m = 51^\circ 1'$ and Δ exceeds π by $46^\circ 38'$. One can easily show that the latter angle is just 45° when $i = 47^\circ 33'$ and when $i = 55^\circ 28'$. Two total reflections at either of these angles of incidence produce a phase difference of 90° between the R_p and R_s components. This phenomenon is applied in the Fresnel rhomb, which may be used like a quarter-wave plate in producing or analyzing circularly and elliptically polarized light.

There is a wave which penetrates a short distance into the second medium when total reflection occurs. It can be proved, however, that there is no energy flux away from the surface. The penetration may be demonstrated by bringing a second surface very close to the totally reflecting surface. The reflection then ceases to be complete, and there is some flux into the second medium across the narrow gap.

PROBLEMS

1. Compute the reflectivity of evaporated aluminum from the values of the indices of absorption and refraction for a wave length of 5461 \AA : $\kappa_0 = 3.5$; $n_0 = 1.15$.

2. Compute the phase shift on perpendicular reflection by a surface of evaporated aluminum for a wave length of 5461 \AA .

3. What should be the angle of incidence such that two total reflections inside crown glass would result in a 90° phase difference between the s and p components if the index of refraction of the glass is 1.5263 ?

4. The indices of refraction and absorption for steel are respectively 2.46 and 1.36 for a wave length of 5893 \AA . What are the values of the principal angle of incidence and the principal azimuth?

² The value of Δ may be reduced by π , to agree with other texts, by reversing the positive direction of R_p in Fig. 12.2.

5. (a) Compute the phase change on normal reflection at a steel surface, using the data given in problem 4. (b) In using Johannson gauge blocks, one observes interference fringes in the wedge-shaped air space between the upper surface of the steel block and the lower surface of a test plate of quartz. At what thicknesses of air film will the interference minima occur in reflected sodium light?

III. RADIATION AND SPECTRA

Chapter 16. Thermal Radiation

16.1 Emissivity and Specific Intensity

Defining **emissivity** e as the emitted radiant flux in watts per square centimeter, and **absorptivity** A as the fraction of the incident flux that is absorbed, Kirchhoff showed that the ratio e/A is a constant for thermal radiators. If this were not true, one could not maintain a uniform temperature between bodies in a uniformly heated, isolated enclosure. Thus for several thermally radiating materials

$$\frac{e_1}{A_1} = \frac{e_2}{A_2} = \frac{e_3}{A_3} = E$$

where E is evidently the emissivity of a perfect absorber or black body for which $A = 1$. Kirchhoff's law shows how the emissivity of any thermal radiator may be obtained from the emissivity of a black body at the same temperature, which turns out to be the most efficient thermal radiator. The laws which apply to the radiation by a black body will be discussed in this chapter.

Specific intensity i is defined as the radiant flux per unit area and unit solid angle. It is similar to the quantity B discussed in Chapter 6, where it was shown that the total flux dF from an element of area $d\sigma$ into a cone of half aperture U is given by

$$dF = \pi i d\sigma \sin^2 U$$

If one considers the emission into an entire hemisphere, $U = \pi/2$, one finds that the emissivity $e = dF/d\sigma$ is related to specific intensity by the formula

$$e = \pi i$$

16.2 Black-body Radiation

In practice, a black body is usually a uniformly heated enclosure with a relatively small aperture through which the radia-

tion escapes to the analyzing and measuring apparatus. The radiation inside the enclosure must be independent of the direction of propagation in regard to its energy, spectral composition, and polarization. This can be shown by introducing a perfect mirror, grating, or polarizer,^c none of which may upset the temperature equilibrium, if no work is done, without violating the second law of thermodynamics.

The **energy density** of radiation u is related to the emissivity e and to the specific intensity i . Consider an arbitrary element of area $d\sigma$ inside a black-body enclosure. The radiant energy passing through it normally into a cone element $d\Omega$ is

$$i \, d\Omega \, d\sigma$$

If the narrow cone is inclined at an angle θ with the normal to the element of surface

$$dF = i \cos \theta \, d\Omega \, d\sigma \quad (1)$$

The energy density u is composed of the energy in all such pencils of rays, the radiant flux per unit area normal to each pencil being given by

$$\frac{dF}{d\sigma} = uc \frac{d\Omega}{4\pi}$$

where c is the speed of light. If the flux is oblique, then

$$dF = uc \cos \theta \frac{d\Omega}{4\pi} d\sigma$$

Since this must be identical with equation (1), it is seen that

$$u = \frac{4\pi i}{c} = \frac{4e}{c}$$

giving the relation among the three quantities usually employed in discussions of radiation.

Thermal radiant energy is distributed continuously throughout a theoretically infinite range of wave lengths. The distribution function for emissivity e_λ is such that the flux between any two wave lengths λ_1 and λ_2 may be found by integration

$$e_{\lambda_1\lambda_2} = \int_{\lambda_1}^{\lambda_2} e_\lambda \, d\lambda$$

The quantity e_λ is thus the radiant flux per unit area and unit wave length increment. Similar distribution functions may be defined for u and i and are related by the equations

$$u_\lambda = \frac{4\pi i_\lambda}{c} = \frac{4e_\lambda}{c}$$

or, for frequency distributions,

$$u_\nu = \frac{4\pi i_\nu}{c} = \frac{4e_\nu}{c}$$

The relation between the two kinds of distribution functions is found as follows. Because

$$u_\nu = \frac{du}{d\nu} = \frac{du}{d\lambda} \frac{d\lambda}{d\nu} \quad \text{and} \quad u_\lambda = \frac{du}{d\lambda}$$

then, since

$$\lambda = \frac{c}{\nu}$$

it follows that

$$|u_\nu| = \frac{c}{\nu^2} u_\lambda \quad \text{and} \quad u_\lambda = \frac{c}{\lambda^2} |u_\nu|$$

The general principles of thermodynamics enable one to derive the general forms of the functions which describe the thermal radiation from a black body. However, one must turn to quantum theory to obtain the exact distribution function and to derive expressions for the radiation constants in terms of fundamental physical constants. Planck was the first to carry out the complete solution of these problems, as will be shown presently.

The law for the total rate of radiation of all frequencies was first given empirically by Stefan and later derived by Boltzmann. One makes use of the first law of thermodynamics and the expression for the radiation pressure, $p = u/3$. The Stefan-Boltzmann law states that

$$u = aT^4$$

In its application it is restricted to black-body radiation. In this case, the experimental value for the constant a in the above formula is

$$a = 7.652 \times 10^{-15} \text{ erg/cm}^3 \text{ deg}^4$$

The theoretical value will be derived in a separate discussion.

The radiation from a black body is distributed continuously throughout the spectrum. By means of thermodynamic arguments, Wien was able to show that the radiation between the limits λ and $\lambda + d\lambda$ is displaced by a change in temperature in accordance with the law

$$\lambda T = \text{constant} \quad \text{or} \quad \frac{T}{\nu} = \text{constant}$$

At the same time, the energy density is changed proportionately to λ^{-5} . Thus, for example, the wave length λ_m of maximum energy density changes in accordance with

$$\lambda_m T = b$$

where b is a constant whose empirical value is 0.2897 cm deg. At the same time, the value of the energy density changes by a factor λ_m^{-5} or as T^5 , since $\lambda_m T$ is constant. The two effects may be combined in a general displacement law of the form

$$u_\lambda = \lambda^{-5} F(\lambda T)$$

or

$$u_\lambda = T^5 f(\lambda T)$$

where $f(\lambda T)$ and $F(\lambda T)$ are different functions of the product λT . One may write the distribution law in terms of frequencies, its form becoming

$$u_\nu = \nu^3 \phi\left(\frac{T}{\nu}\right)$$

where $\phi(T/\nu)$ is a function of the quotient T/ν . Neither the constants nor the forms of these functions can be determined by means of thermodynamics alone. One may, however, test the indicated form of the distribution law by plotting u_λ/T^5 against λT for black-body radiation. The points are found to fall on one curve, Fig. 16.1, regardless of the temperature; whereas if one just plots u_λ against λ , he obtains a different curve for each temperature T .

Max Planck in 1900 applied the quantum theory to derive his famous distribution formula, from which the formulas discussed above may also be obtained. He imagined a uniformly heated enclosure filled with "oscillators" in equilibrium with the radiation. He assumed finite elements of energy ϵ distributed at random among the oscillators. Only the number of quanta of energy per

oscillator is of importance. The number of ways W by which a distribution may be realized is proportional to the probability of the distribution. Representing the number of quanta by P and the number of oscillators by N , one may divide the P quanta into N groups by means of $N - 1$ partitions to achieve one possible distribution thus:

$$\epsilon \epsilon | \epsilon | \epsilon \epsilon \epsilon | \epsilon \epsilon \epsilon \epsilon | \epsilon \text{ etc.}$$

The number of ways in which these $N + P - 1$ objects can be arranged is $(N + P - 1)!$, but of these, $(N - 1)!P!$ are identical.

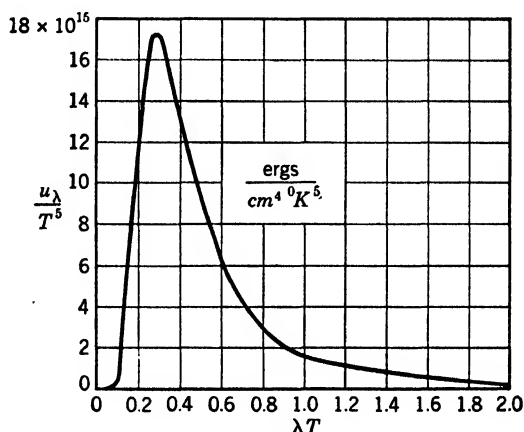


FIG. 16.1. Distribution curve from which the radiation density u_λ in a black body may be determined for any temperature T and at any wave length λ .

Thus the total number of different ways in which one may distribute the P quanta among the N oscillators is

$$W = \frac{(N + P - 1)!}{(N - 1)!P!}$$

which is proportional to the probability of this particular system P, N .

Let S_N denote the entropy of the system. Boltzmann has shown that entropy is related to probability as follows:

$$S_N = k \log_e W$$

The basis for this equation is fundamentally the fact that the entropy of a combination of systems, or parts of a system, is the

sum of the entropies of each, while the total probability is the product of the probabilities of the parts. By combining the above two equations, one finds that

$$S_N = k \{ \log (N + P - 1)! - \log (N - 1)! - \log P! \}$$

Using Stirling's approximation for large factorials, namely,

$$\log N! \doteq N \log N - N$$

the remaining terms in Stirling's formula being relatively small when N is very large, it is found that

$$S_N = k \{ (N + P - 1) \log (N + P - 1) - (N - 1) \log (N - 1) - P \log P \}$$

Let \bar{U} be the average energy per oscillator, so that

$$\bar{U} = \frac{P\epsilon}{N} \quad \text{or} \quad P = N \frac{\bar{U}}{\epsilon}$$

Eliminating P by the use of this formula and neglecting unity in the first logarithm because it is relatively small when P and N are very large, it is found that

$$S_N = k \left\{ \left(N + N \frac{\bar{U}}{\epsilon} - 1 \right) \log N \left(1 + \frac{\bar{U}}{\epsilon} \right) - (N - 1) \log (N - 1) - N \frac{\bar{U}}{\epsilon} \log N \frac{\bar{U}}{\epsilon} \right\}$$

The average entropy per oscillator is then

$$S = \frac{S_N}{N} = k \left\{ \left(1 + \frac{\bar{U}}{\epsilon} \right) \log \left(1 + \frac{\bar{U}}{\epsilon} \right) - \frac{\bar{U}}{\epsilon} \log \frac{\bar{U}}{\epsilon} \right\}$$

According to the second law of thermodynamics, the absolute temperature T is derivable from the entropy by the operation:

$$\frac{\partial S}{\partial \bar{U}} = \frac{1}{T}$$

Hence

$$\frac{1}{T} = k \left\{ \frac{1}{\epsilon} \log \left(1 + \frac{\bar{U}}{\epsilon} \right) - \frac{1}{\epsilon} \log \frac{\bar{U}}{\epsilon} \right\}$$

and therefore

$$\frac{\epsilon}{kT} = \log \frac{1 + (\bar{U}/\epsilon)}{(\bar{U}/\epsilon)}.$$

Solving this for the average energy \bar{U} of an oscillator, one finds that

$$\bar{U} = \frac{\epsilon}{e^{\epsilon/kT} - 1}$$

The energy density u_ν of the radiation in equilibrium with the oscillators is related to \bar{U} by the equation

$$u_\nu = \frac{8\pi\nu^2}{c^3} \bar{U}$$

This may be derived in various ways, for example, by computing the radiation from a Hertzian oscillator by the use of classical electrodynamics, as shown by Page in *Theoretical Physics*, pages 479-481.

By combining the last two equations, one finds that

$$u_\nu = \frac{8\pi\nu^2}{c^3} \frac{\epsilon}{e^{\epsilon/kT} - 1}$$

According to Wien's displacement law, the correct formula for u_ν must be of the form

$$u_\nu = \nu^3 \phi\left(\frac{T}{\nu}\right)$$

It is readily seen that Planck's formula will have the required form if the quantum of energy ϵ is proportional to the frequency. Accordingly, one writes

$$\epsilon = h\nu$$

and Planck's radiation formula becomes

$$u_\nu = \frac{8\pi h\nu^3}{c^3} (e^{h\nu/kT} - 1)^{-1}$$

The constant h is appropriately called Planck's constant. Its first appearance in physical theory was in Planck's development of the equation for the energy distribution of thermal radiation. It is

evident that the other distribution functions for specific intensity or emissivity can be readily obtained from u_ν .

The Wien displacement formula, $\lambda_m T = \text{constant}$, may be shown to follow from Planck's law by changing u_ν to u_λ and differentiating¹ with respect to λ to find the condition for a maximum in u_λ . The resulting formula is

$$\lambda_m T = \frac{ch}{4.9651k}$$

in which the constant quantity on the right is expressed in terms of the fundamental physical constants h , c , and k . They may be measured by entirely independent methods, providing a check of the radiation theory.

Similarly, one may also obtain the Stefan-Boltzmann law by integrating² Planck's distribution law to obtain the total radiant energy of all wave lengths or frequencies. In this way it can be shown that

$$u = \left(\frac{8\pi^5 k^4}{15c^3 h^3} \right) T^4$$

One may compute values of h and k from the experimental values of the constants in the last two formulas, using Michelson's value for the speed of light c . The determinations are in good agreement with standard values, although this is not a particularly accurate method for determining Planck's constant h and Boltzmann's constant k . Moreover, one may obtain a value for Avogadro's constant N_0 by dividing k into the universal gas constant R per mole and may then obtain the value of the electronic charge e by dividing N_0 into Faraday's constant of electrolysis. The results are astonishingly good.

The Rayleigh-Jeans classical distribution formula is obtained from Planck's when $h\nu/kT$ is so small that one may represent the exponential in the denominator by the first two terms of its series expansion as follows:

$$e^{h\nu/kT} \doteq 1 + \frac{h\nu}{kT}$$

¹ RICHTMYER and KENNARD, *Introduction to Modern Physics*, McGraw-Hill, 4 ed., 1947, page 181.

² RICHTMYER and KENNARD, *Introduction to Modern Physics*, McGraw-Hill, 4 ed., 1947, page 182.

One then finds that

$$u_\nu = \frac{8\pi\nu^2}{c^3} kT$$

This formula, which may be used at long wave lengths or low frequencies, was first derived by Rayleigh and by Jeans, using purely classical, non-quantum theory. At the other extreme, when the exponential term is very large compared with unity, one obtains **Wien's distribution formula**

$$u_\nu = \frac{8\pi h\nu^3}{c^3} e^{-h\nu/kT}$$

This is a useful approximation in the visible region if the temperature is not too high, as is true of most incandescent light sources.

16.3 Selective Thermal Radiation

If the absorption coefficient is denoted by A , the emissivity of any thermal radiator will be A times that of a black body, in accordance with Kirchhoff's law. For a constant value of A , the intensity is reduced without a change in the form of the distribution curve or the location of the wave length at which the intensity is greatest. The radiator is then referred to as a **gray body**.

For metals, however, the absorption coefficient often varies as the square root of the absolute temperature divided by the wave length, while in other cases it may vary in a complicated manner. It follows that the form of the distribution curve and the magnitude of the total emission are altered. In general, the total emission no longer varies as the fourth power of the absolute temperature; for metals, it usually varies as the fifth power. Some materials, such as oxides of the rare earths, have sharp absorption bands and consequently show a relatively high emission at these wave lengths. Selective thermal radiation is often utilized in the construction of incandescent sources of light. For example, because of such selectivity, the tungsten filament and the Welsbach mantle radiate a larger portion of their radiant energy in the visible region.

16.4 Color Temperature

The color temperature of a body is that temperature of a black body which has the same ratio of specific emissions as the selective radiator in two intervals, λ_1 to $\lambda_1 + d\lambda_1$ and λ_2 to $\lambda_2 + d\lambda_2$. In general, the color temperature may have various values according to the choice of the two wave lengths. However, if λ_1 and λ_2 are in the visible region, the color temperature is often quite unambiguous and may be loosely defined as the temperature of a black body which has the same color as the selective radiator.

For several metals, $\log A_\lambda$ may be expressed as a linear function of $1/\lambda$. One may then write Wien's distribution law in the form

$$E = c'e^{\alpha/\lambda}c_1\lambda^{-5}e^{-c_2/\lambda T} = c'c_1\lambda^{-5}e^{-(c_2-\alpha T)/\lambda T}$$

where T is the actual absolute temperature of the selective radiator. For a black body at the temperature T_c , the corresponding distribution formula is

$$E_B = c_1\lambda^{-5}e^{-c_2/\lambda T_c}$$

The ratio of the emissivities E/E_B will be the same at all wave lengths, or any pair whatsoever, if

$$\frac{c_2 - \alpha T}{\lambda T} = \frac{c_2}{\lambda T_c}$$

or

$$\frac{1}{T_c} = \frac{1}{T} - \frac{\alpha}{c_2}$$

By definition, T_c is the color temperature of the selective radiator whose actual temperature is T . The quantity T_c is perfectly unambiguous if the absorptivity follows the law assumed above, in other words, if

$$A_\lambda = c'e^{\alpha/\lambda}$$

For a gray body A_λ is independent of wave length, and consequently α is zero. Then it is seen that $T_c = T$, indicating that the color temperature of a gray body is equal to its true temperature. Usually, α is positive, if the law holds even approximately, and the color temperature is greater than the true temperature. A few

examples of the color temperatures of some well-known sources are as follows:

SOURCE	COLOR TEMPERATURE (T_c), °K
Hefner flame	1840
Normal carbon filament	2080
Normal Mazda lamp	2380
Normal nitrogen-filled lamp	2600–2900
"Daylight lamp" (incand.)	4500
No. 1 photoflood	3480
No. 2 photoflood	3430
100-watt household lamp	2830
Mean sunlight	5600

PROBLEMS

1. Planck's distribution law is often simplified in the visible region by dropping the 1 in the denominator, which gives Wien's distribution law. For what temperatures will the computed intensities be accurate to 0.1 per cent?

2. The absorption coefficient of a metallic surface radiator is approximately proportional to the square root of the quotient of the absolute temperature divided by the wave length. What will be the algebraic form of the Planck distribution formula for this kind of selective thermal radiation? How will the distribution compare with that for a black body?

3. If the intensity at wave length 0.665μ in the spectrum of a black body at the melting point of platinum, 1750°C , is 100 arbitrary units, what is the temperature of a black body giving just half this intensity? This problem suggests how a "radiation scale of temperature" may be established.

4. (a) Compute the energy density per unit wave length interval at a wave length of 1μ for a black body at 2000°K . (b) What is the specific intensity of the radiation coming from a small hole in a uniformly heated enclosure at this temperature and wave length?

5. From the experimental values 0.2897 cm deg for Wien's constant, and $7.652 \times 10^{-15} \text{ erg/cm}^3 \text{ deg}^4$ for Stefan-Boltzmann's constant, compute the values of Planck's constant h and Boltzmann's constant k .

Chapter 17. Line and Band Spectra

17.1 Discontinuous Spectra

The emission and absorption spectra of gases and vapors are discontinuous, with the exception of certain broad continua associated with molecular and solid-state energy levels. They may be excited thermally by means of a flame or by an electric furnace, preferably operated under reduced pressure. However, one usually employs an electric glow discharge, an arc, or a spark.

Radiating atoms or ions give rise to line spectra, whereas radiating molecules are characterized by **band spectra**. The latter consist of many lines which are so close together that they are not usually resolved by a single-prism spectroscope. Because the lines have a non-uniform distribution in frequency and intensity, they frequently appear as groups of broad bands with sharp edges or "band heads" which may be either on the high- or on the low-frequency side. A good grating is needed to study the structure of band spectra.

Although **line spectra** usually show no obvious evidence of orderliness in structure, they are highly characteristic of the emitter and the physical conditions under which they are excited. This is stated concisely in the law of Bunsen and Kirchhoff (1859) as follows: "Under given conditions, each element emits a perfectly definite spectrum which is characteristic of that element alone." Applications of this law have led to the discovery of new elements, to spectrochemical analysis in industrial laboratories, and to the study of atomic structure as well as the processes occurring in light sources, including the stars.

In astronomical spectroscopy one frequently encounters the **Doppler-Fizeau effect**, which is the change in the observed frequency or wave length of light when there is relative motion of the source with respect to the observer. For example, consider the source to be approaching the observer with a velocity component v . The number of waves ν emitted in one second will be contained

in an interval $c - v$. Thus the wave length in the line of sight will be

$$\lambda' = \frac{c - v}{\nu}$$

The observed frequency is the number of these waves passing the observer per second, their velocity relative to him being c . Then

$$\nu' = \frac{c}{\lambda'} = \nu \left(\frac{c}{c - v} \right)$$

ν being the frequency at the source, or at the observer in the absence of relative motion.

This effect has been used to measure the velocity of streams of ionized gas in the laboratory as well as in the solar atmosphere. The velocities of many stars relative to the earth have also been determined by the Doppler effect. An interesting observation is that the more distant stars and galaxies are in general moving more rapidly away from the earth than those less distant.

17.2 Spectral Series

The first evidence for orderly structure in line spectra was found in 1885 by Balmer in the spectrum of atomic hydrogen. Balmer observed that the wave lengths of the lines in the visible and near ultraviolet spectrum of a glow discharge in hydrogen could be represented by the relatively simple formula

$$\lambda = 3645.6 \frac{n^2}{n^2 - 4}$$

the number n taking on successive integral values starting with three, the resulting wave length being in angstrom units. Since the advent of the quantum theory, it has been preferable to write this formula in terms of frequencies ν or wave numbers $\bar{\nu}$. Thus

$$\bar{\nu} = 109\,677 \left(\frac{1}{2^2} - \frac{1}{n^2} \right); \quad n = 3, 4, 5, \dots$$

gives the number of waves per centimeter in vacuo. The lines related by a formula of this general character are said to be mem-

bers of a **spectral series**. The names of some of the spectral series of hydrogen and their formulas are:

$$\text{Lyman series: } \bar{\nu} = 109,677 \left(\frac{1}{1^2} - \frac{1}{n^2} \right); \quad n = 2, 3 \dots$$

$$\text{Balmer series: } \bar{\nu} = 109,677 \left(\frac{1}{2^2} - \frac{1}{n^2} \right); \quad n = 3, 4 \dots$$

$$\text{Paschen series: } \bar{\nu} = 109,677 \left(\frac{1}{3^2} - \frac{1}{n^2} \right); \quad n = 4, 5 \dots$$

etc., for the Brackett series and the Pfund series.

As the value of n increases, the intensity of the lines diminishes. Usually one can observe only the first four or five members of any series, although as many as forty have been measured in the Balmer series. Figure 17.1 shows the general appearance of a spectral series.

The spectra of hydrogen and of hydrogen-like atoms may be most readily deduced by an application of Bohr's theory. These atoms have the simple structure of a nucleus and a single planetary electron. They include hydrogen, singly ionized helium, doubly ionized lithium, etc. Let Ze and M represent the charge and mass

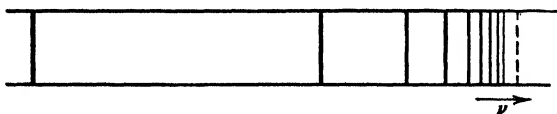


FIG. 17.1. A typical spectral series.

of the nucleus, and let $-e$ and m represent the corresponding quantities for the single electron revolving in an orbit at a distance a from the nucleus. For simplicity, the orbit is assumed to be circular. The potential energy of the system is

$$-\frac{Ze^2}{a}$$

and the kinetic energy is $\frac{1}{2}I\omega^2$, where I is the moment of inertia about the center of mass, and ω is the angular velocity. Writing the kinetic energy in terms of the individual masses and their

separation, one finds that

$$\begin{aligned}\frac{1}{2} I \omega^2 &= \frac{1}{2} \left[m \left(\frac{M}{M+m} \right)^2 + M \left(\frac{m}{M+m} \right)^2 \right] a^2 \omega^2 \\ &= \frac{1}{2} \left(\frac{mM}{M+m} \right) a^2 \omega^2\end{aligned}$$

Since the electrostatic attraction is equal to the centripetal force,

$$\frac{Ze^2}{a^2} = m\omega^2 r$$

or

$$\frac{Ze^2}{a^2} = m\omega^2 \frac{M}{M+m} a = \left(\frac{mM}{M+m} \right) \omega^2 a$$

Substituting this result into the expression for the kinetic energy, one finds that

$$\frac{1}{2} I \omega^2 = \frac{1}{2} \frac{Ze^2}{a}$$

Thus the sum of the kinetic and potential energy is given by

$$E = -\frac{1}{2} \frac{Ze^2}{a}$$

Since atomic spectra are discrete, Bohr inferred that the energy must also have only certain allowed values. The orbits were accordingly restricted to those having an angular momentum equal to an integral multiple of $h/2\pi$, where h is Planck's constant. Thus

$$I\omega = n \frac{h}{2\pi}$$

or

$$\left(\frac{mM}{M+m} \right) a^2 \omega = n \frac{h}{2\pi}$$

Squaring both sides of this equation and dividing them respectively by the two sides of the centripetal force equation to eliminate ω^2 , one finds that

$$\left(\frac{Mm}{M+m}\right)a^3 = \frac{n^2 h^2 a^2}{4\pi^2 Ze^2}$$

which leads to

$$a = n^2 \frac{(M+m)h^2}{4\pi^2 MmZe^2}$$

giving the radii of the allowed orbits. The expression for the energy of the n th state is accordingly

$$E_n = -\frac{Ze^2}{2a} = -\frac{2\pi^2 mc^4 Z^2}{\left(1 + \frac{m}{M}\right) h^2 n^2}$$

Bohr further assumed that these states are non-radiating, but that during a transition from any state to one of lower energy a single quantum is emitted in accordance with

$$h\nu = E_{n_2} - E_{n_1}$$

where ν is the frequency of the radiation. In terms of wave numbers this reduces to

$$\tilde{\nu} = \frac{E_{n_2}}{hc} - \frac{E_{n_1}}{hc}$$

One may think of this expression as a difference between two **spectroscopic terms** in wave-number units. Letting

$$R_\infty \equiv \frac{2\pi^2 mc^4}{h^3 c} = 109,737.42 \text{ cm}^{-1}$$

these term values T_n reduce in this case to

$$T_n \equiv -\frac{E_n}{hc} = \frac{R_\infty}{1 + (m/M)} \frac{Z^2}{n^2}$$

The wave number of a spectral line arising from a transition from an energy level n_2 to a level n_1 can be computed from

$$\tilde{\nu} = T_{n_1} - T_{n_2} = \frac{R_\infty Z^2}{1 + (m/M)} \left(\frac{1}{n_1^2} - \frac{1}{n_2^2} \right)$$

For hydrogen, since $Z = 1$ and $M/m = 1836.6$, this reduces to the empirical formula already given, the Balmer series being obtained

if $n_1 = 2$ and $n_2 = 3, 4, 5, \dots$. For singly ionized helium, $Z = 2$, one obtains the Pickering series when $n_1 = 4$ and $n_2 = 5, 6, 7, 8, \dots$. It is easy to see that every other line of this series would coincide with a Balmer series line if it were not for the term m/M in the denominator. Even the isotopes of hydrogen give slightly different wave lengths because of the finite and different values of the nuclear masses M for each isotope.

Shortly after Balmer's discovery, Rydberg showed that there are spectral series in the spectra of other elements and that these may be represented by a formula similar to that given above, namely,

$$\tilde{\nu} = \frac{R}{(n_1 + \sigma_1)^2} - \frac{R}{(n_2 + \sigma_2)^2}$$

or, simply

$$\tilde{\nu} = T_1 - T_2$$

In these formulas R is the **Rydberg constant** and is essentially the same for the arc spectra of all elements. There is, however, a small change due to nuclear mass, as explained above. For any one series, T_1 is a **fixed term** representing the convergence limit, and T_2 is the **running term**, which takes on successively decreasing values as n_2 increases by steps of unity. The quantity σ is a quantum correction which characterizes each series of terms. It is not entirely independent of n .

In the case of the alkalis and often other elements, one finds several superimposed series, of which the most important are designated as the principal P , sharp S , diffuse D , and fundamental F series. They are distinguished by different values of σ ; thus one may write the principal series as

$$\tilde{\nu}_P = T_{\infty P} - \frac{R}{(n - \sigma_P)^2}$$

where $n = 2, 3, 4, \dots$ for Li, $n = 3, 4, 5, \dots$ for Na, etc. For the sharp series one writes similarly

$$\tilde{\nu}_S = T_{\infty S} - \frac{R}{(n - \sigma_S)^2}$$

It has been found that $T_{\infty P}$ is equal to the greatest value of the running term for the sharp series. Thus, for sodium, where $n = 3$ for this term,

$$T_{\infty P} = \frac{R}{(3 - \sigma_S)^2}$$

which may be written symbolically as $3S$. The running terms are specified by the indefinite number n , those for the principal series being designated as nP terms. For sodium, $n_1 = 3$, the various series may be designated by

$$\tilde{\nu}_P = 3S - nP$$

$$\tilde{\nu}_S = 3P - nS$$

$$\tilde{\nu}_D = 3P - nD$$

$$\tilde{\nu}_F = 3D - nF$$

The indicated relations between the fixed and running terms of the various series are in accordance with empirical laws.

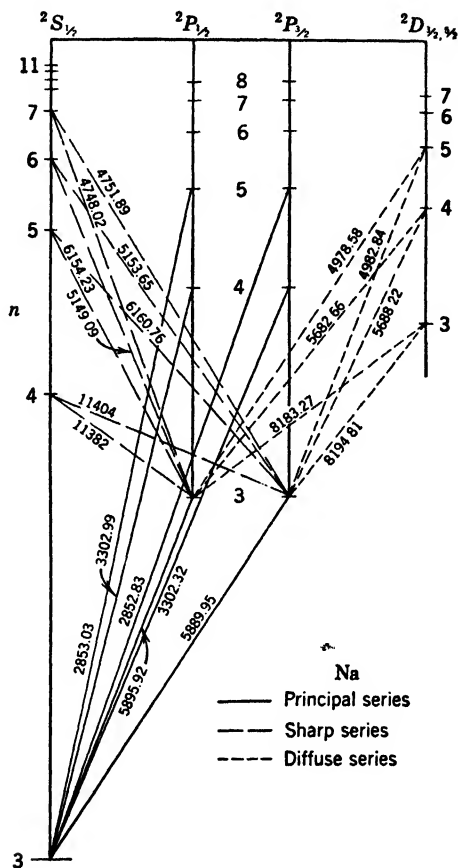
In order to account for the **fine structure** of the above terms, the quantum number J was introduced by Sommerfeld. It indicates the magnitude of the total angular momentum of the atom in Bohr units of $h/2\pi$. This "total" is in most cases merely the vector sum of the orbital angular momentum L and the electron spin momentum S . The above term symbols S , P , D , F , G , etc., respectively indicate that L has the values 0, 1, 2, 3, 4, etc. The value of L for one valence electron is equal to the orbital quantum number l for that electron; but if there is more than one valence electron, L is any one of the various vector sums of l which differ by unity from zero to the maximum possible. The same may be said of S , the total spin, which is obtained from $s = \frac{1}{2}$ for each individual electron by vector addition. For each possible J value there is a corresponding sublevel of the spectral term. The total number of these sublevels is $2S + 1$ if $L > S$, and $2L + 1$ if $L < S$. It is customary, in either case, to call $2S + 1$ the **multiplicity** of the term and to indicate it by a superscript to the left of the term symbol. The J value is written as a subscript. Thus for the D lines of sodium one may write

$$\tilde{\nu} = 3^2S_{\frac{1}{2}} - 3^2P_{\frac{1}{2}} \quad \text{for } \lambda \text{ 5896 } \text{\AA}$$

$$\tilde{\nu} = 3^2S_{\frac{1}{2}} - 3^2P_{\frac{3}{2}} \quad \text{for } \lambda \text{ 5890 } \text{\AA}$$

Other lines have similar designations. The entire spectrum may be conveniently summarized by means of transitions be-

tween the levels of a **term diagram** or **energy-level diagram**, Fig. 17.2, each spectral term corresponding to an energy level in the theory of atomic structure. There are certain **selection**



to separate the terms into columns of different L values. Then any one term can combine generally with terms in adjacent columns ($\Delta L = \pm 1$) subject to the indicated limitations on the values of ΔJ . Downward transitions correspond to emission,

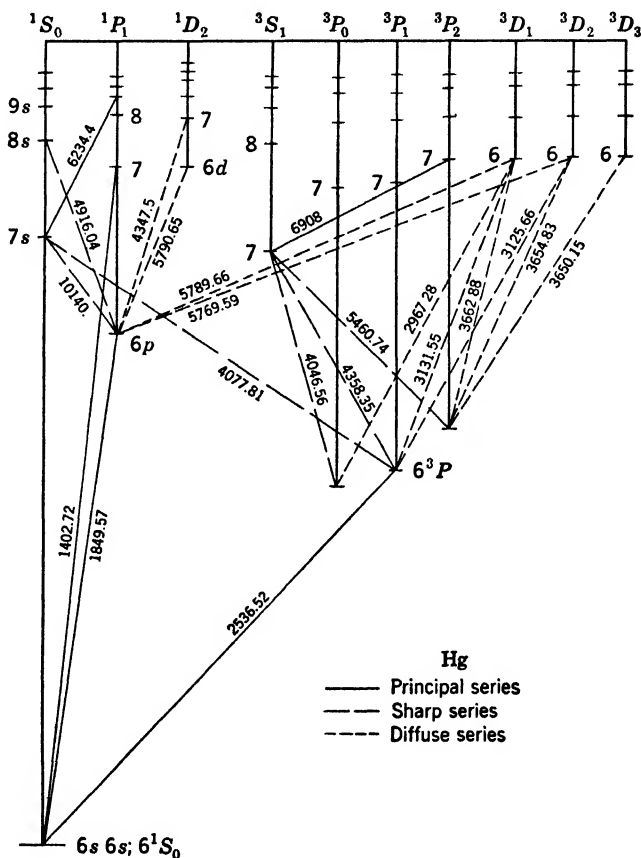


FIG. 17.3. Energy-level diagram for the mercury atom, showing some of the transitions commonly observed.

whereas upward transitions, for which the same selection rules hold, correspond to absorption frequencies. At ordinary temperatures, the absorption will be almost exclusively from the ground, or lowest, state because of the enormous preponderance of atoms in this state. Thus the observed absorption spectrum,

consisting mainly of the principal series, is simpler in structure than the emission spectrum.

Energy-level diagrams for sodium and mercury are shown in Figs. 17.2 and 17.3, with the transitions ordinarily observed in the laboratory indicated. Because mercury has two valence electrons, S may have the values zero or unity. The former value gives rise to the **singlet states**, whereas the latter gives the **triplet states**. Exact values of the spectroscopic terms may be found in Bacher and Goudsmit's *Atomic Energy States*.

17.3 Wave Mechanics

Although the energy levels of hydrogen-like atoms may be computed accurately by Bohr's theory, it has been found impossible to develop a satisfactory extension of the theory to more complicated systems. Even in regard to the hydrogen problem, Bohr's theory is not entirely satisfactory because (a) it introduces the quantization of angular momentum by a perfectly arbitrary assumption, (b) it does not give a method for computing relative intensities of lines, (c) it gives an incorrect fine-structure pattern.

Three different forms of a new theory have been developed since 1926, principally by Schrödinger, Heisenberg, and Dirac. Schrödinger's **wave mechanics** lends itself most readily to a brief description. It is based on the concept introduced by De Broglie, that material particles are associated with waves having a wave length given by

$$\lambda = \frac{h}{mv}$$

called the **De Broglie wave length**, where m is the mass of the particle, v is its velocity, and h is Planck's constant. This formula has been verified directly by measurements of diffraction angles of electron beams using crystals as gratings.

In wave mechanics, the quantized states of an atom correspond to the various possible stationary De Broglie wave patterns, each being characterized by integral numbers of nodes. For example, corresponding to a simple circular Bohr orbit in the old theory, there is now a stationary wave with an integral number n of waves in a distance equal to the circumference $2\pi r$ of the orbit. Thus

$$n \frac{h}{mv} = 2\pi r$$

This leads to the equation for angular momentum

$$mvr = n \frac{h}{2\pi}$$

which Bohr arbitrarily adopted in specifying the allowed orbits.

The above is only an enormously simplified illustration, but it serves to reveal a physical foundation for Bohr's quantum condition. In general, one must proceed to solve Schrödinger's equation, which is a differential equation for waves in three dimensions, and then adjust the constants in the solutions to fit appropriate boundary conditions. Acceptable solutions must be single valued, finite, and continuous, and must vanish at infinity. Three integers automatically appear in each solution in the role of quantum numbers. They are related to the numbers of various kinds of nodal surfaces in the three-dimensional stationary vibration. The square of the amplitude of the wave gives the probability of finding the electron at any location. One may alternatively compute the probable momentum, but not at any definite location, so that orbits, as continuous trajectories, cannot be determined. Hence one now speaks of **wave functions** instead of orbits, the wave functions being acceptable solutions of Schrödinger's equation. In some problems in atomic physics, it is convenient to consider the orbital electrons as an "electron cloud" about the nucleus, the density of the cloud at each point being proportional to the square of the amplitude of the wave function. Examples of such clouds for the hydrogen atom are illustrated by White in the *Physical Review*, **37**, 1417 (1931).

The energy values required in Schrödinger's equation in order to obtain each acceptable wave function give the sequences of energy levels observed in spectroscopy. Transitions between levels can be proved to give the observed frequencies. One may also compute the transition probabilities and thus obtain the relative intensities of spectral lines and derive the selection rules for transitions. The results previously derived by an application of Bohr's theory to hydrogen-like atoms are checked exactly, but one may proceed much farther with the new theory. For a discussion of the theory and applications of quantum mechanics, the reader is referred to Pauling and Wilson's *Introduction to Quantum Mechanics*.

17.4 Spectrochemical Analysis

Spectrochemical analysis is based on the law of Kirchhoff and Bunsen, Art. 17.1. While studying the spectra of the alkalis, they were able to discover and isolate two new elements, rubidium and cesium, by means of spectral lines. They also investigated the solar spectrum and explained the **Fraunhofer lines** as due to the absorption by various elements in the atmosphere of the sun or of the earth. At least fifty-nine elements have now been identified in the solar atmosphere.

For the purpose of spectrochemical analysis, the most suitable source often depends on the character of the material to be analyzed. Conducting solids may be used as electrodes in an electric arc or spark. If only a small sample is available, or if the material is non-conducting, it may be necessary to pack it into holes drilled into rods of pure graphite, silver, or copper. Of course, blank tests with the electrodes will have to be run in order to determine which lines are due to the electrodes themselves. In this test it must be noted that the altered discharge, after the sample is inserted, may cause the intensification, or even appearance, of some weak lines in the electrode spectrum. Solutions or liquid samples may be placed in a cup which serves as a lower electrode, or they may be dried on the end of a pure graphite rod. Sometimes an atomized spray of solution is injected into a flame, as in the method developed by Lundegard. Gases are conveniently analyzed by the use of an electrical discharge in a tube at low pressure. If high-frequency current is used, the electrodes may be outside the tube, thus avoiding possible contamination of the gas, as in Experiment 11.

The temperatures of most sources dissociate molecules, so that only the line spectra of the elements are observed. This is usually advantageous. On the other hand, a flame or a gaseous discharge tube will often give band spectra due to molecules and parts of molecules.

In order to detect small amounts of any element, one must examine the spectrum for the sensitive lines. They are often in the ultraviolet. Table 17.1 by Hartley shows what is meant by a sensitive or **persistent line**. The source was a spark to the surface of a solution containing cadmium chloride. Evidently wave lengths 3260.2 and 2321.6 Å, in the central part of the table, disappear between 1 and 0.1 per cent cadmium, whereas λ 2265.8 can be observed even with concentrations of 0.001 per cent and

somewhat lower. Similar persistent lines in the spectra of various elements are given in many spectroscopic tables and are usually listed by element as well as by wave length. Their presence or absence in a spectrum should be ascertained before one concludes that any particular element is present or not. Their absence,

TABLE 17.1

WAVE LENGTHS OBSERVED AT LOW CONCENTRATIONS OF CADMIUM

Cd: 1%	0.1%	0.01%	0.001%
3609.6 Å	3609.6 Å	3609.6 Å
3466.7	3466.7	3466.7
3465.2	3465.2	3465.2
3402.8	3402.8
3260.2
2747.7	2747.7
2572.3	2572.3
2321.6
2313.5	2313.5	2313.5
2288.8	2288.8	2288.8
2265.8	2265.8	2265.8	2265.8
2196.4
2146.8

however, may merely mean that the element is not abundant enough to detect with the source, spectrograph, and photographic technique used. This possibility must be tested by the use of chemically similar samples of known concentration.

Tables such as the M.I.T. tables of wave lengths list about 100,000 lines between 2000 Å and 10,000 Å or, on the average, 12 lines per angstrom unit. Hence the measurement of the wave length of a line to 1 Å is far from sufficient for unambiguous identification. However, a knowledge of the probable composition of the source often helps exclude some interfering lines. It is advisable to look for typical groups or combinations of lines with characteristic separations and intensity distributions. The greater the number of lines that one can find with the proper relative intensities, the surer one may be of the presence of any given element. A single line would require a measurement to an accuracy of better than 0.05 Å for positive identification, and such accuracy is often impossible to attain.

The sensitivity of trace detection varies greatly. In general, metals can be detected more readily than non-metals. In the case

of the alkalis it is not particularly difficult to detect concentrations of only one part in a million. On the other hand, it is interesting that von Tongeren reports a sensitivity of only 0.3 per cent tin in cassiterite, the chief ore of tin. Sawyer is of the opinion that this low sensitivity is due to the use of visual lines instead of the more sensitive ultraviolet lines. In general, the accuracy of spectrochemical detection of traces is superior to that obtained by chemical wet methods and as good as or better than the colorimetric methods. The speed is generally greater if suitable charts are available.

The basis of quantitative spectrochemical analysis is generally some modification of the homologous line-pair method of Gerlach and Schweitzer. **Homologous lines** are lines whose intensities vary in approximately the same manner with discharge conditions. One of the lines of the pair is due to the element being measured, while the other is a neighboring line due to some element present in equal amounts in the various samples. The latter may be a line of the matrix or of some element added to each sample in equal amounts. The vapor pressures of both elements and the excitation functions of both lines should be similar. Many such line pairs may be found listed in the literature. As an example a portion of one of Gerlach and Schweitzer's tables is given (Table 17.2). By the use of such tables and a visual examination of the

TABLE 17.2
ANALYSIS OF CADMIUM IN TIN

Wave Lengths of Lines, Å	Intensities Equal at % by Wt. of Cd	Remarks
Cd λ 3404 Sn λ 3341	10%	Give short exposure; fixed point sharp; extremely invariant.
Cd λ 3404 Sn λ 3219	0.5%	Give long exposure; fixed point sharp; extremely invariant.
Cd λ 3466 Sn λ 3219	0.15%	Give long exposure; sharply defined; very invariant.
Cd λ 2288 Sn λ 2282	0.01%	Extremely long exposure; somewhat in- definite fixed point; very invariant.

lines, one may determine concentrations with an accuracy of about 25 per cent. To obtain greater accuracy it is necessary to measure the relative intensities of some line pair for a series of known

samples and for the unknown sample. One must use a suitable light source, good photographic technique, and an accurate method for measuring photographic densities. The densities should be

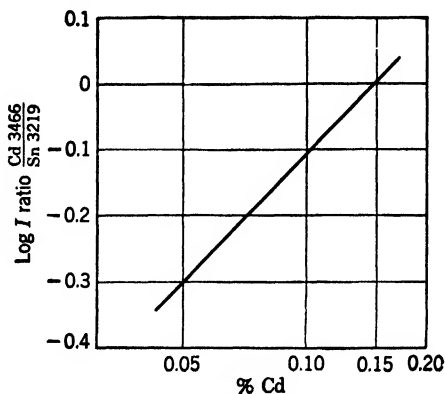


Fig. 17.4. Example of a calibration curve for quantitative spectrochemical analysis.

between about 0.3 and 1.0 for the best results. An example of an analytical curve is shown in Fig. 17.4. To measure the intensity ratio, one must first apply a suitable method of plate calibration

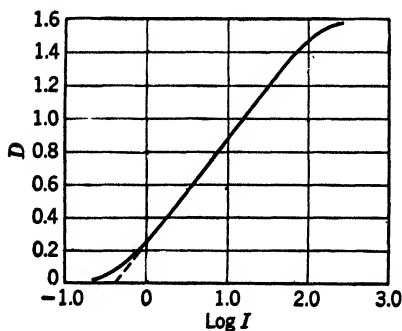


Fig. 17.5. The characteristic curve of a photographic plate.

to establish the variation of density with the intensity of the spectral line. A typical characteristic curve giving this relation is shown in Fig. 17.5. The logarithm of the intensity ratio for a line pair is readily found from such a curve by the measurement of the densities of the two lines being compared. To obtain a curve of this kind, one must impress on the plate a suitable series of calibration marks of known

relative intensity. These may be obtained in various ways, as explained below.

The **reciprocity law** states that photographic density is a function of the product of the intensity and the time of exposure.

If one could show that the reciprocity law holds for a given emulsion, he might calibrate the plate merely by using a sequence of increasing exposure times. However, this law is extremely unreliable at low and at high intensities, the exact range of its validity depending on the emulsion. One therefore obtains more trustworthy results by varying the intensity. This can be done (1) by applying the inverse square law; (2) by a stepped slit, after Duffendack; (3) by the use of calibrated wire screens, after Harrison; (4) by a neutral absorbing wedge, after Follet; or (5) by the use of close groups of spectral lines of known relative intensity, after Vincent and Sawyer.

A popular calibration technique employs a rotating stepped sector or logarithmic sector which varies the average intensity logarithmically along the length of the slit. Such a rotating screen really varies the total time of exposure by the use of intermittent flashes of the same frequency but of different duration. It has been shown empirically, however, that, if the rate of rotation of the sector is above some critical value, the photographic density obtained is the same as for a continuous exposure to the time average of the intensity. Webb is of the opinion that the frequency must be high enough so that no more than one quantum of radiant energy strikes a given grain of silver halide during each flash. Sawyer states that for some fast emulsions no critical frequency can be found. At any rate, it appears that this method of plate calibration should be tested for the emulsion used, before relying on it.

Instead of plotting densities against log intensities, as in the conventional characteristic curve, it is more convenient to plot galvanometer deflections against intensities, using log-log or semi-log paper. It is preferable to "reverse" the galvanometer scale so that the higher readings correspond to the greater densities (less transmitted light). A plate calibration curve of this kind is similar in form to the characteristic curve shown in Fig. 17.5. The logarithm of the ratio of the intensities of any two analysis lines may be readily obtained from the two corresponding galvanometer deflections.

The light source used for quantitative work must be selected with care if one is to obtain reproducible results. One popular light source is the a-c arc developed by Duffendack.¹ The trans-

¹ DUFFENDACK, *Industrial and Engineering Chemistry*, **10**, 161 (1938).

former used is a conventional pole-type transformer used for power distribution in cities. The input voltage of 110 or 220 volts is stepped up to about 2300 volts and supplied to the arc gap between graphite electrodes through a ballast resistance to limit the current to about 2.5 amperes. Solutions are dried on the ends of the graphite electrodes with the aid of an infrared drying lamp. Another popular source is the Fuessner spark, for which the circuit diagram is given in Fig. 17.6. A voltage of about 15 to 20 kilovolts is applied across a condenser of 0.003 to 0.2 microfarad, which discharges through an inductance of 0.01 to 0.05 milli-

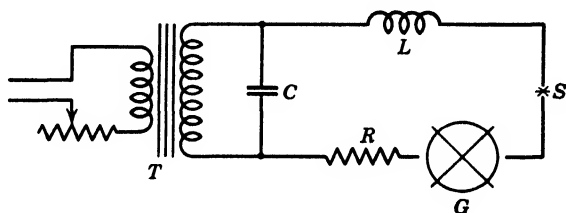


FIG. 17.6. Circuit diagram of a Fuessner spark source S . T , high-voltage transformer; C , condenser; L , inductance; R , resistance; G , rotary spark gap.

henry, a rotary spark gap, the source gap, and a damping resistance. The rotary gap is operated by a synchronous motor, and the phase of the contacts is adjusted so that a train of sparks passes across the source gap at the peak of each a-c cycle. The oscillations are damped by the resistance and by the use of an air blast. The capacitance of the condenser and the kva rating of the transformer must be selected so that the condenser becomes practically fully charged in each interval between breakdown of the gap.

The spectrograph should have a sufficiently high dispersion to separate the lines to be measured with a slit width of 20 to 50 μ . The wide slit is used to obtain a flat-topped line suitable for densitometry. The background of scattered and reflected light should be low because it is difficult to make accurate corrections for any appreciable background.

17.5 Band Spectra

The dense groups of lines often observed in the spectra of arcs arise from the radiation of molecules. In a spectroscope of low resolving power, the structure is not resolved and one often finds

what appears to be a group of continuous bands which usually are sharply bounded on either the violet or the red side. These bands sometimes occur in a large number of groups. With a spectroscope of high resolving power, one observes that the intensity distribution in a band is not continuous, but that each band consists of many lines which are very close together at the band head. The variable spacing of the lines, as well as their intensities, accounts for their general appearance as continuous bands at low resolutions. Molecular spectra of simpler structure than those in the visible and in the ultraviolet are found in the near infrared and in the far infrared regions. Although the complexity of molecular spectra naturally increases with that of the molecule, many of the characteristics of these spectra are shown by considering a diatomic polar molecule such as one of the hydrogen halides.

As before, we will consider the frequency radiated or absorbed to be proportional to the difference between the energy of two stationary states. The kinds of energy which must be considered in establishing molecular energy levels are: (1) rotational, (2) vibrational, and (3) electronic; the third being due to changes in electron configuration, as in atomic spectra.

17.6 Rotational Bands

Consider a dipole, Fig. 17.7, rotating about a principal axis normal to the line joining the nuclei of the two atoms and passing through the center of mass, C.M. According to Bohr's theory, the angular momentum will be an integral multiple of $h/2\pi$, that is

$$I\omega = K \frac{h}{2\pi}; \quad K = 0, 1, 2, 3, \dots$$

where I is the moment of inertia, ω the angular velocity, and h Planck's constant. The moment of inertia is

$$I = \frac{m_1 m_2}{m_1 + m_2} r^2$$

where m_1 and m_2 are the masses of the atoms, and r is the separation of their centers of mass. The rotational energy is then

$$E_K = \frac{1}{2} I \omega^2 = K^2 \frac{h^2}{8\pi^2 I}$$

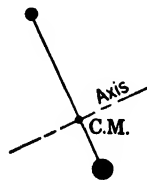


FIG. 17.7. A diatomic molecule rotates about an axis through its center of mass.

Wave mechanics leads to a similar result except that one obtains the factor $K(K + 1)$ instead of K^2 . Thus the correct expression is

$$E_K = K(K + 1) \frac{h^2}{8\pi^2 I}$$

To show the similarity of this expression to the classical result except for a "zero-point" energy and half-integral quantum numbers, one may easily transform the above expression algebraically into the form

$$E_K = \left(K + \frac{1}{2}\right)^2 \frac{h^2}{8\pi^2 I} - \frac{1}{4} \frac{h^2}{8\pi^2 I}$$

Either expression for E_K defines a system of energy levels between which downward transitions indicate emission, and upward transitions indicate absorption. As usual, the observed transitions are restricted by selection rules. For spontaneous emission

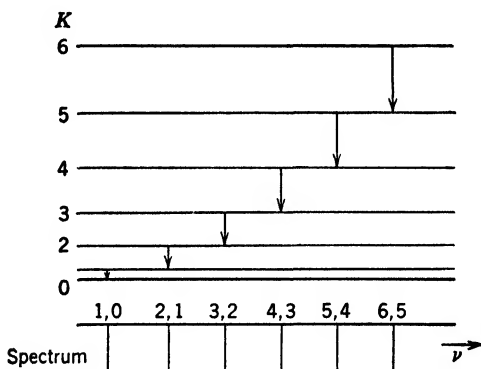


FIG. 17.8. Rotational energy levels and allowed transitions. The resulting spectrum is shown below the vertical arrows.

of light, K decreases by unity, while for absorption, K increases by unity. An energy-level diagram, indicating some transitions resulting in emission, is shown in Fig. 17.8, with the associated spectrum given below it. The emission frequencies are given by

$$\nu_{K+1,K} = \frac{h}{8\pi^2 I} [(K + 1)(K + 2) - K(K + 1)] = (K + 1) \frac{h}{4\pi^2 I}$$

Thus the lines should have a uniform spacing of

$$\Delta\nu = \frac{h}{4\pi^2 I}$$

Experimentally, one finds that this is not quite the case, presumably due to the increase in I at higher rotational frequencies, which causes a stretching of the molecular bond. Thus Czerny finds experimentally that the wave numbers in the rotational spectrum of hydrogen chloride are given by

$$\bar{\nu} = 20.794M - 0.00164M^3; \quad M = 1, 2, 3 \dots$$

These bands are in the extreme infrared.

When M is small, one may neglect the second term, obtaining a spacing of

$$\Delta\bar{\nu} = \frac{h}{4\pi^2 I c} = 20.794 \text{ cm}^{-1}$$

This gives a value for I of $2.66 \times 10^{-40} \text{ gm cm}^2$, or a separation between nuclei of $1.28 \times 10^{-8} \text{ cm}$, which is of the correct order of magnitude.

17.7 Vibration-rotation Bands

At greater molecular energies, the atoms in the molecule may vibrate along the line joining their nuclei with the accompanying exchanges of kinetic and potential energy typical of simple harmonic motion. The potential energy of the molecule varies with the distance between nuclei in the manner shown by the heavy line in Fig. 17.9. The equilibrium separation corresponding to a minimum potential energy is denoted by r_0 . Near this value of r the curve approximates a parabola, $E = \frac{1}{2}kr^2$, so that the restoring force is $F = kr$ and the vibrations are practically harmonic. Quantum mechanics leads to values of vibrational energy E_v given by

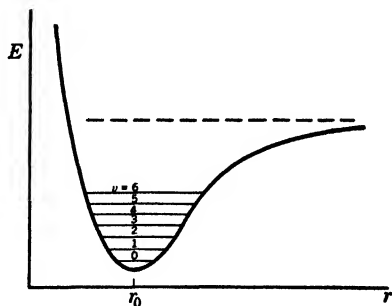


FIG. 17.9. Potential energy curve with seven vibrational levels indicated.

$$E_v = (v + \frac{1}{2})h\nu_0; \quad v = 0, 1, 2, 3, \dots$$

where ν_0 is the classical frequency,

$$\nu_0 = \frac{1}{2\pi} \sqrt{\frac{k}{m}}$$

m being the mass and k the force constant. The vibrational energy levels are represented in Fig. 17.9 by the horizontal lines.

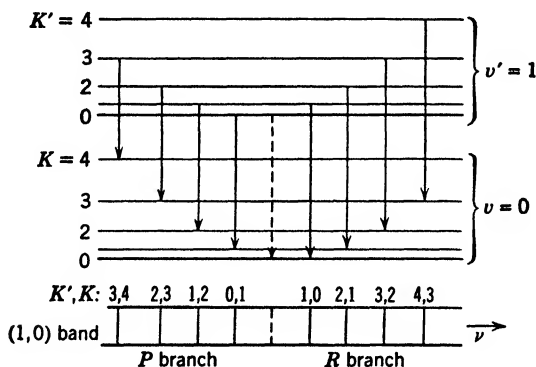


FIG. 17.10. Energy-level diagram for a vibration-rotation band. The solid lines with arrows pointing downward indicate allowed transitions. The resulting spectrum is shown below.

Associated with each vibrational energy, the molecule may have various amounts of rotational energy, as was explained in Art. 17.6. If one neglects interaction between vibration and rotation, the combined energy is

$$E = \left(v + \frac{1}{2}\right)h\nu_0 + \frac{K(K+1)h^2}{8\pi^2I}$$

For a harmonic oscillator, v may change only by unity. Actually, one may also find bands due to greater changes in v , but they are generally quite weak. Figure 17.10 shows a portion of an energy-level diagram for vibration and rotation with some of the allowed transitions indicated. The frequencies of spectral lines emitted or absorbed are given by

$$\nu = \frac{E' - E''}{h} = (v' - v'')\nu_0 + \frac{h}{8\pi^2I} (\pm 2K'' \pm 1 + 1)$$

where the upper signs apply if K decreases on emission (R branch) and the lower, negative, signs are used if K increases (P branch). The intensities of the members in both branches rise to a maximum as K'' increases, the location of the maximum being farther from the center of the band at higher temperatures. Such bands, having the general intensity distribution and structure shown in Fig. 17.11, appear in the near infrared. If the resolution of the spectroscope is too low to show the rotational structure, one observes only the two peaks, which respectively correspond to the envelopes of

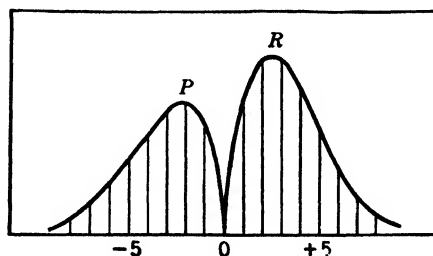


FIG. 17.11. Variation of intensity of the lines in the P and R branches of a vibration-rotation band.

the P and R branches. The separation of neighboring lines in the rotational fine structure is approximately that given for the rotational bands, except that the innermost lines are twice as far apart.

Experiment shows the expected agreement between the line separations for low values of K in the rotation and in the vibration-rotation bands. In the latter case, the value for $\Delta\bar{\nu}$ in the spectrum of hydrogen chloride is found to be 20.8 cm^{-1} , which is in excellent agreement with the rotational value given before. As K increases, the value of I is no longer constant and the separations are found to be unequal.

17.8 Electronic Bands

Electronic energy levels of molecules are analogous to the energy levels of atoms. Various configurations of electrons, or combinations of wave functions, occur which are classified primarily according to the components of electronic angular momentum along the polar axis of the molecule. Thus the symbols Σ , π , and Δ represent energy states in which this component of angular momentum is respectively 0, 1, and 2 Bohr units. Furthermore, these

states may be singlet, doublet, or triplet, according to whether the electron spin resultant is 0, $\frac{1}{2}$, or 1 Bohr unit. For further details of classification, the reader is referred to books dealing with molecular spectra. Because of the large energy difference between the electronic levels, the spectra resulting from transitions between them are generally in the visible or ultraviolet regions. If all three forms of energy, electronic, vibrational, and rotational, are considered, the frequency condition becomes

$$h\nu = h\nu_e + (E_v' - E_v'') + (E_r' - E_r'')$$

where $h\nu_e$ is the difference between the energies of the electronic levels, and the changes in vibrational energy and in rotational energy are respectively denoted by $(E_v' - E_v'')$ and $(E_r' - E_r'')$. In the rotational term, it must be noted that the moments of inertia are usually quite different in the initial and final states, since these correspond to different electron configurations. Figure 17.12 shows part of the vibrational-rotational structure of two electronic levels, as well as some of the transitions between the rotational states associated with $v' = 1$ and $v'' = 0$. Substituting the expressions for vibrational and rotational energy into the frequency formula leads one to the following expression for the wave numbers of a single band:

$$\bar{\nu} = \bar{\nu}_e + \frac{v'\nu_0' - v''\nu_0''}{c} + \frac{h}{8\pi^2c} \left[\frac{K'(K' + 1)}{I'} - \frac{K''(K'' + 1)}{I''} \right]$$

The selection rule for K is $\Delta K = 0$ or ± 1 . Each of these three possibilities gives rise to a different "branch." If we let

$$\bar{\nu}_e + \frac{v'\nu_0' - v''\nu_0''}{c} \equiv A$$

$$\frac{h}{16\pi^2c} \left(\frac{1}{I'} + \frac{1}{I''} \right) \equiv B$$

$$\frac{h}{8\pi^2c} \left(\frac{1}{I'} - \frac{1}{I''} \right) \equiv C$$

these branches may be written as follows in terms of $K = K''$:

1. *P* branch: $K' = K'' - 1 = K - 1$

$$\bar{\nu}_P = A - 2BK + CK^2; \quad K = 1, 2, 3, \dots$$

2. *Q* branch: $K' = K'' = K$

$$\tilde{\nu}_Q = A + CK + CK^2; \quad K = 0, 1, 2, \dots$$

3. *R* branch: $K' = K'' + 1 = K + 1$

$$\tilde{\nu}_R = A + 2B(K + 1) + C(K + 1)^2; \quad K = 0, 1, 2, \dots$$

Each equation represents a parabola, or rather a portion of one, since K is always positive. Either the *P* or the *R* branch turns back over the band, depending on the sign of C , which depends on the

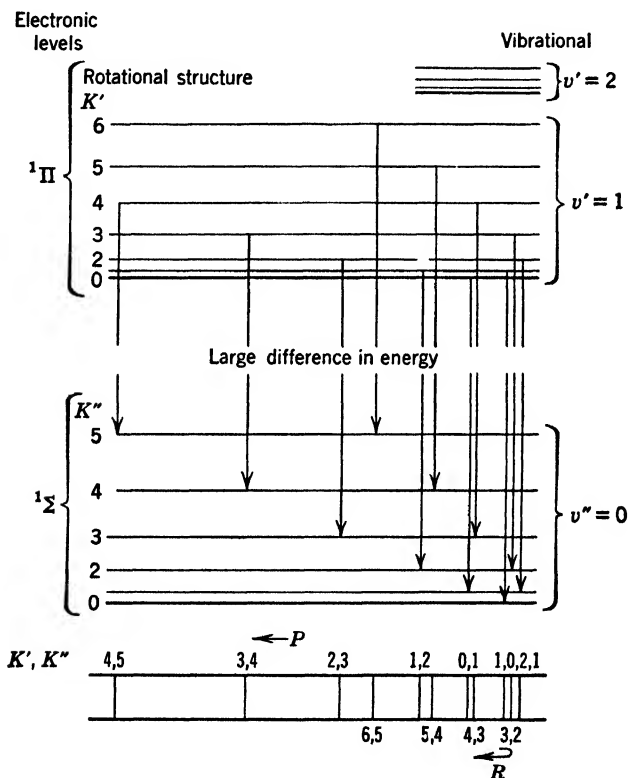


FIG. 17.12. Energy-level diagram for an electronic band. The band structure is indicated below the vertical arrows.

relative magnitudes of I' and I'' . The band shown in Fig. 17.12 is for a *negative* C , or I' greater than I'' . The band head is at the "blue" edge of the band, which is said to be **degraded toward**

the red. It will be noted that, if $C = 0$, the Q branch is a single line. The Q branch does not appear at all if the total angular momentum around the internuclear axis is zero for both states, for then $\Delta K = 0$ becomes a "forbidden transition." Consequently, there is no Q branch in electronic bands resulting from $\Sigma - \Sigma$ transitions. Some bands show a doublet structure as in Fig. 17.13, which is a band in the absorption spectrum of oxygen.

The energy levels of liquids and solids are broadened because of the proximity of the molecules. This is in accordance with a form

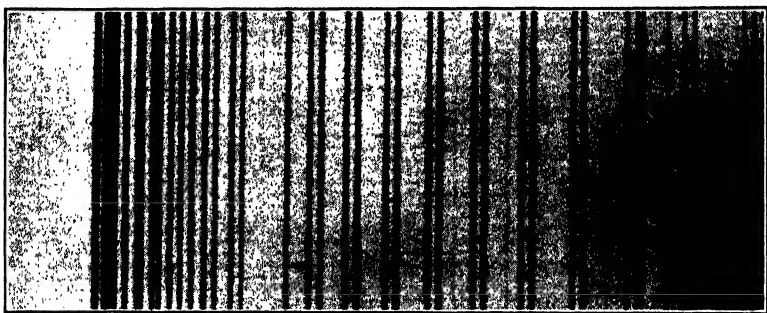


FIG. 17.13. Spectrogram of an oxygen absorption band in the red region of the solar spectrum. The band head at the left, wave length 6870 Å, is the Fraunhofer "line" B. (Photographed by H. A. Rowland.)

of Heisenberg's principle, which states that a short lifetime, or small uncertainty in duration of a state, is accompanied by a correspondingly large uncertainty in energy, i.e., $\Delta E \Delta t \sim h$. There may also be a displacement of a level due to the change in potential energy arising from the charges on surrounding atoms. Consequently, the absorption bands of solids and liquids are generally very broad with no perceptible fine structure, in contrast with those due to the same material in the vapor state. Compounds of the rare earths are exceptional in having relatively sharp absorption lines. They arise from transitions between filled levels and the unfilled 4f subshell, which is buried under completed 5s, 5p, and 6s subshells and is therefore quite free from external influences.

17.9 Microwave Absorption

Pure rotational spectra of molecules sometimes extend into the microwave region, where accurate measurements may be made of

differences between closely spaced energy levels. It is also relatively easy to measure the absolute absorption coefficients from which transition probabilities may be determined. The spectra of most molecules are quite simple in the **microwave region**, where the wave lengths are of the order of magnitude of one centimeter. Because of the simple spectra, the transitions may be identified without great difficulty. Water vapor, for example, has a strong absorption band at 0.7418 cm^{-1} or at a wave length of 1.35 cm. The absorption here amounts to about 4.7 per cent per kilometer and gram of water vapor per cubic meter. This absorption is of importance in radar and has consequently been studied rather thoroughly.²

The absorption by water vapor is fundamentally due to its dipole moment. Since oxygen is electrically non-polar, one would expect, at first sight, no absorption of microwaves by dry, un-ionized molecules of oxygen. The oxygen molecule, however, is paramagnetic, having a permanent magnetic dipole moment. This permits an absorption due to upward transitions induced by microwaves between states of different magnetic moment. This absorption should be weaker than the electrical dipole absorption because the Bohr magneton is smaller than the Debye unit of electrical polarity. Thus it is only in extremely long paths that the effect should be observable. The absorption maximum is computed to lie at a wave length of about 0.5 cm, and the amount of absorption due to the oxygen in dry air at this wave length, at 300°K and an air pressure of 76 cm of mercury, is predicted to be 27.6 per cent per kilometer. This falls to less than 2.5 per cent at wave lengths above 0.56 cm and at wave lengths below 0.45 cm. Of course, the absorptivity varies exponentially with distance, there being a constant percentage reduction in intensity in each unit distance.

The equivalent frequency differences $\Delta W/h$ between hyperfine structure components and between Zeeman effect components of an energy state, Art. 19.4, are frequently located in the microwave region. By subjecting atoms or molecules to magnetic fields of the equivalent frequencies, one may often induce transitions between these spectroscopically nearby levels. The transitions which may be induced are subject to the usual rules for allowed changes in the

² VAN VLECK, *Physical Review*, **71**, 425 (1947); KING, HAINER, and CROSS, *Physical Review*, **71**, 433 (1947).

quantum numbers. Instead of observing absorption, one may detect the occurrence of transitions by utilizing the associated change in magnetic moment of the particles in a narrow beam passing through oppositely directed, non-homogeneous magnetic fields at *A* and *B*, Fig. 17.14. The beam which is originally projected through narrow slits along the solid curve in Fig. 17.14(*b*) will be deflected, as shown by one or the other of the dotted lines, when the magnetic moment has been changed at *C*. In this event, the beam will miss the exit slit at *D* and there will be a reduction in the

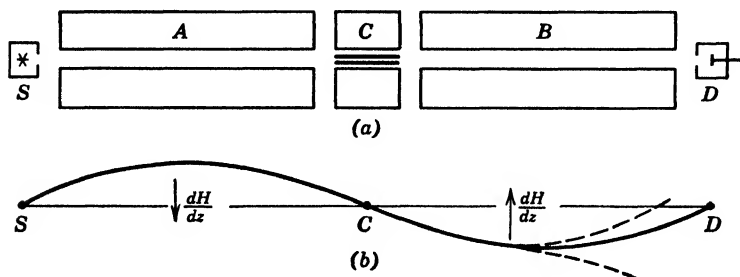


FIG. 17.14. (*a*) Molecular beam apparatus. The beam from the source *S* leaving at an angle with *SCD* is deflected in the direction of dH/dz in the non-uniform magnetic field *A*, then in the opposite direction in the field *B*, being detected at *D*. The path is shown exaggerated in (*b*).

intensity of the beam. The transitions are induced by a steady field *H* at *C* and a high-frequency field at right angles to *H*. The frequencies at which such changes occur are then interpreted in terms of differences between energy levels. The steady field *H* causes a Zeeman splitting of the energy states (Art. 19.4) and a Larmor precession of the molecular magnets about the direction of the field. Transitions are observed when the oscillating field has the same frequency as that corresponding to the difference between two magnetic levels, this being, classically, the frequency of the Larmor precession.

For example, the fine structure of the second quantum state of hydrogen consists of $2^2P_{1/2}$, $2^2S_{1/2}$, $2^2P_{3/2}$ states. It is now thought that the second state is only 0.033 cm^{-1} above the first, while the third is 0.332 cm^{-1} above the second. Ordinary spectroscopical measurements of these intervals are very difficult because of the fact that the separations are very small fractions of the wave number of the line being examined. Thus even 0.332 cm^{-1} represents

a change of only 2.2 thousandths of a per cent of the wave number of the hydrogen alpha line, which is $15,233.01 \text{ cm}^{-1}$. Moreover, the third quantum state, which is also involved in the emission of this line, is itself complex, making the structure of the alpha line quite complicated. On the other hand, 0.332 cm^{-1} corresponds to a frequency of about 10,000 megacycles per second, or a wave length of 3 cm. Since there has been a great advance in this "radar" region during the late war, it is now possible to produce and measure such a frequency with an accuracy of about 0.01 per cent. Thus the atomic-beam method enables one to measure the interval $2^2S_{1/2} - 2^2P_{3/2}$, in this example, to a far greater accuracy than by observing the structure of the alpha line with the best of interferometer techniques. The new method is limited to components of normal states or to meta-stable states, and to intervals which correspond to allowed transitions.

The application of a steady magnetic field to the atomic beam at C produces observable Zeeman splitting of energy states with fields as low as 0.05 gauss. An atom in a magnetic field H has its energy changed by the amount

$$W = -\mu_z H$$

where μ_z is the component of the magnetic moment of the atom in the direction of the field. According to quantum mechanics, only specified values of μ_z may exist, corresponding to the magnetic quantum numbers m which specify the components of angular momentum in the direction of H , which differ by increments of $h/2\pi$. A high-frequency field of just the right frequency is capable of inducing transitions from one state of space-quantized angular momentum to another, subject to the condition that $\Delta m = \pm 1$.

In the hydrogen problem discussed above, the frequencies actually³ observed corresponded to transitions between one of the Zeeman components of $2^2S_{1/2}$ and any component of either $2^2P_{1/2}$ or $2^2P_{3/2}$. The relative locations of the parent levels were then found by extrapolating toward zero field, where the components merged.

The molecular-beam method has been used by Rabi and his associates in measuring the magnetic moments and angular momenta of nuclei. In a constant field H at C , Fig. 17.14, a particle

³ LAMB and RETHERFORD, *Physical Review*, **72**, 241-243 (1947).

with a magnetic moment $g\mu_0$ precesses about the field with the Larmor frequency

$$\nu = g\mu_0 \frac{H}{h}$$

where μ_0 is the Bohr unit of magnetic moment, $eh/4\pi Mc$. If now an oscillating magnetic field of frequency ν is simultaneously applied, transitions to another level of different magnetic quantum number may be induced. Thus, by measuring the resonance frequency ν and the field H , one may compute the gyromagnetic ratio g for the particle. Even when one has a coupled system having a rotational magnetic moment as well as a nuclear moment, he may use a field of sufficiently high value so that the moments are decoupled and the two types of transitions occur independently. It is then possible to measure one nuclear magnetic moment in the presence of other moments. The procedures are outlined in a review article by Kellogg and Millman in the *Reviews of Modern Physics*, **18**, 323-352 (1946), in which numerous references are given to the original literature, and some of the results are tabulated.

PROBLEMS

1. Compute the wave numbers and the wave lengths of the first lines and the limits of the Lyman series, the Balmer series, and the Paschen series of hydrogen. $R = 109,677 \text{ cm}^{-1}$.
2. Compute the difference between (a) the wave numbers and (b) the wave lengths of the hydrogen and of the deuterium alpha lines.
3. Compute a "spectroscopic value" for the specific charge e/m of the electron from the Rydberg constants of hydrogen (109,677.76) and of helium (109,722.40). The atomic weights are respectively 1.008123 and 4.00390, and the Faraday constant is 9648.9 emu/gram-atomic weight of univalent ion.
4. Compute the series limit and the quantum defect for the mercury $6^3P_0 - m^3S_1$ series, assuming a Rydberg formula and using the first two lines whose wave lengths are 5460.74 Å and 3341.48 Å in air at 15°C and 760 mm pressure. What is the accepted value of the series limit? Explain the discrepancy.

Chapter 18. X-ray Spectra

18.1 Characteristic X-ray Emission and Absorption

Whereas optical spectra originate in the transition of one of the valence, or outermost, electrons from one energy state to another, x-ray line spectra originate in transitions involving an electron in one of the inner groups. Principally as a result of the study of x-ray absorption (and emission) spectra, the grouping of structural electrons illustrated in Table 18.1 has been determined. Only the elements to be considered in this discussion are included.

TABLE 18.1
EXAMPLES OF ELECTRON CONFIGURATIONS

n		1	2	3	4	5	6
l		0	0 1	0 1 2	0 1 2 3	0 1 2	0
Atomic number							
1	H	1					
2	He	2					
3	Li	2	1				
10	Ne	2	2 6				
11	Na	2	2 6	1			
16	S	2	2 6	2 4			
17	Cl	2	2 6	2 5			
18	A	2	2 6	2 6			
19	K	2	2 6	2 6	1		
29	Cu	2	2 6	2 6 10	1		
42	Mo	2	2 6	2 6 10	2 6 5	1	
74	W	2	2 6	2 6 10	2 6 10 14	2 6 4	2
79	Au	2	2 6	2 6 10	2 6 10 14	2 6 10	1

The levels to the left of the staircase are fully occupied. Consequently they remain unchanged in the succeeding elements. As a result of this, there is a great similarity between the x-ray spectra of the various elements, except that as one examines spectra of the lighter elements one group of lines after another disappears with the electron shell necessary for its emission. The *K*-series lines result from a transition to a vacancy in the *K* level, while *L*-series lines result from a transition to a vacancy in the *L* level, the initial vacancies being created by the bombardment of the

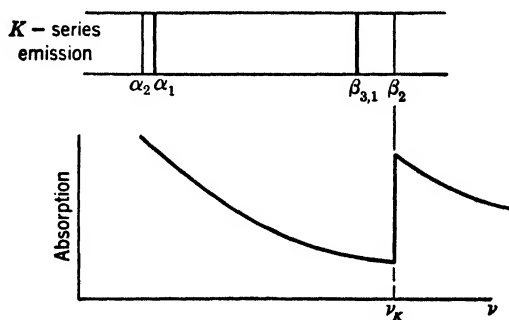


FIG. 18.1. The structure and relative location of the *K* characteristic emission lines and the *K* absorption edge.

target by high-speed electrons or by quanta of high energy. In the early spectrograms the *K* series was observed to consist of only two lines, α and β , while the *L* series apparently consisted of three, α , β , and γ . However, these were soon resolved into the components shown by the illustrations of emission spectra in Figs. 18.1 and 18.2. Below the emission spectra are shown the corresponding absorption spectra. One can trace any one of the lines or absorption edges over a wide range, from high to low atomic numbers, before it disappears. The variation of frequency with atomic number is approximately expressed by Moseley's law:

$$\sqrt{\nu} \propto Z$$

or, better,

$$\sqrt{\nu} \propto Z - s$$

where s is a screening constant, so that $Z - s$ is the effective nuclear charge. The curves obtained by plotting $\sqrt{\nu}$ against Z or

$Z - s$ are, however, not perfectly straight. Furthermore, lines do not vanish with the associated energy level, persisting for two or three elements beyond, but with large deviations from Moseley's law. These are the so-called "semioptical" lines of Siegbahn and Backlin, in which there is presumably a transition of an electron from an optical level to an x-ray level.

The occurrence of absorption edges in x-ray spectra, as against absorption lines in optical spectra, becomes clear when the character of the levels is considered. For the optical electron, the levels

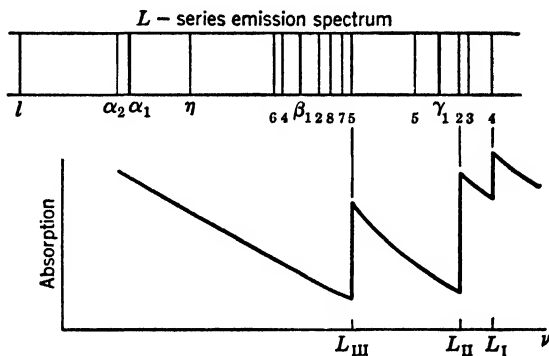


FIG. 18.2. Typical L characteristic emission lines and absorption edges.

of higher energy are all unoccupied, whereas for the x-ray electron they are usually fully occupied. Thus no x-ray absorption occurs until the frequency of the incident radiation becomes high enough to remove an x-ray electron from within the atom by a photoelectric process. Each absorption edge is the threshold for one of the different kinds of photoelectric ionization. It is from a measurement of these photoelectric thresholds that one obtains the number of the groups of electrons and their energies. The x-ray scattering process is the cause of a contribution to absorption which becomes more important at higher frequencies. The classical theory of scattering is presented in Art. 21.2. For the quantum theory of this process, the reader is referred to Compton and Allison's *X-rays in Theory and Experiment*. Still other absorption processes are considered in Art. 19.7.

X-ray absorption is usually expressed in terms of the fractional reduction of intensity I per gram of absorber per unit area. Thus

$$\frac{dI}{I} = -\mu_m dm$$

where m is the mass per unit area, and the coefficient μ_m is called the **mass absorption coefficient**. Integrating this expression, one finds that

$$I = I_0 e^{-\mu_m m}$$

where I_0 is the incident intensity. Between absorption edges, the mass absorption coefficient varies about as the cube of the wave length and the fourth power of the atomic number Z . Thus

$$\mu_m = C\lambda^3 Z^4 + b$$

the quantity b being nearly constant at about 0.2 over a wide range of wave lengths. This term is due to x-ray scattering and is relatively small compared with the first term, except at short wave lengths.

To obtain the characteristic line emission spectra, one of the inner electrons must first be removed from the atom, causing a K_1 ionization or an L_1 ionization, etc. The transition of some outer electron to fill this vacancy with a change in energy ΔW results in the radiation of a spectrum line of frequency

$$\nu = \frac{\Delta W}{h}$$

or wave number

$$\tilde{\nu} = \frac{\Delta W}{hc}$$

The lines which correspond to transitions to a vacancy in the K shell are called K -series lines. The more intense lines arise from the more probable transitions, which are as a rule from the nearest group of levels. In a similar way, the lines of the L , M , etc., series are produced.

18.2 X-ray Energy Levels

The above discussion gives a qualitative survey of the origin of x-ray spectra. The quantitative relations are best summarized by means of an energy-level diagram. It is an interesting fact that the K , L , M , etc., energy levels, when one ignores for the moment their fine structure, follow approximately the scheme of levels of

the simple hydrogen-like atom of atomic number $(Z - s_{nl})$. It appears that one electron subtracted from a more or less closed shell system is equivalent to just one electron associated with a nucleus of appropriate charge. Thus to a first approximation

$$W = Rhc \frac{(Z - s_{nl})^2}{n^2} = h\nu_0 = hc\tilde{\nu}_0$$

$$\frac{W}{Rhc} = \frac{(Z - s_{nl})^2}{n^2} = \frac{\tilde{\nu}_0}{R}$$

where R is Rydberg's constant, $109,737 \text{ cm}^{-1}$; h is Planck's constant, $6.62 \times 10^{-27} \text{ erg sec}$; $n = 1$ for $\tilde{\nu}_K$; $n = 2$ for $\tilde{\nu}_L$; $n = 3$ for $\tilde{\nu}_M$, etc. Moseley's law for the absorption edges follows immediately. The number n is called the **principal quantum number** since it fixes the energy to the first approximation. These energies may be conveniently indicated by an energy-level diagram. In x-ray diagrams it is customary to take the normal state as the zero level. The K level will then lie above it by the amount $(Z - s_1)^2$ in Rydberg-frequency units, which is $1/Rhc$ times the energy required to raise the atom to this level. The L , M , N , etc., levels will lie below the K level with rapidly increasing crowding.

This crowding becomes so great that the diagram, Fig. 18.3, is not drawn to scale in this respect. The levels obtained as the result of a first approximation in which screening is neglected are shown in the left-hand column of the diagram. The energy states corresponding to each value of n may have various amounts of **orbital angular momentum**, which is designated by a second quantum number l . This number may have all possible integral values (including zero) which are at least one unit less than the number n . Since the average distance of an electron from the nucleus for any given n also depends on l , being greater for the higher values of l , the screening of the nucleus by the other electrons will increase as l increases. Consequently, the energy levels are displaced and split into sublevels as shown in the central column of Fig. 18.3. By the use of the theory of relativity as first applied by Sommerfeld, it is found that the energy of each state is increased by the amount

$$\frac{\alpha^2(Z - d_{nl})^4}{n^3} \left(\frac{1}{l + 1/2} - \frac{3}{4n} \right)$$

where

$$\alpha^2 = \frac{2\pi e^2}{hc}$$

is Sommerfeld's **fine structure constant**. Finally, if, following Uhlenbeck and Goudsmit, the electron is assumed to have a spin

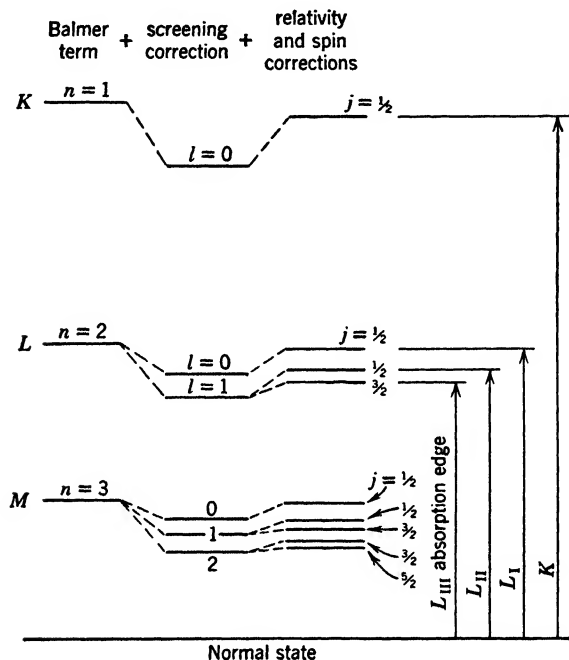


FIG. 18.3. X-ray energy-level diagram, showing fine structure of levels and transitions giving rise to K and L absorption edges.

of $\frac{1}{2}h/2\pi$, which may combine with the angular momentum of the orbit in two opposite senses, the energy levels suffer further modification. Those for which l is not zero are split into two levels. A new quantum number j is introduced which indicates the resultant of the angular momentum of the atom and the spin of the electron. For the case of a single electron, there are two possibilities

$$j = |l - \frac{1}{2}|$$

and

$$j = |l + \frac{1}{2}|$$

The additional energy due to electron spin is

$$-\alpha^2 \frac{(Z - d'_{nlj})^4}{n^3} \frac{j(j+1) - l(l+1) - s(s+1)}{2l(l + \frac{1}{2})(l+1)}$$

The relativity and spin corrections change the energy-level scheme to its final form, which is shown in the third column of Fig. 18.3. This gives us a means of accounting for the results of absorption experiments, some of the processes being indicated by the upward arrows on the right in Fig. 18.3.

When emission spectra are fitted into this energy-level scheme, it is found that only certain of the energy differences ΔW result in observable emission lines. This has led to the formulation of the selection rules

$$\Delta n > 0; \quad \Delta l = \pm 1;$$

$$\Delta j = \pm 1 \text{ or } 0$$

These, however, are not perfectly rigid, since faint lines for which $\Delta l = 2$ have occasionally

been observed. In view of the selection rule for l , it is convenient to separate the energy levels into parallel columns according to the value of this quantum number, as shown in Fig. 18.4. The principal transitions can then be readily indicated by slanting lines between adjacent columns in the figure. The first rule is satisfied by all downward transitions. The third rule eliminates one of the transitions between each pair of doublet levels in adjacent columns. Figures 18.5 and 18.6 are x-ray spectrograms showing the K -series lines of iron and copper respectively.

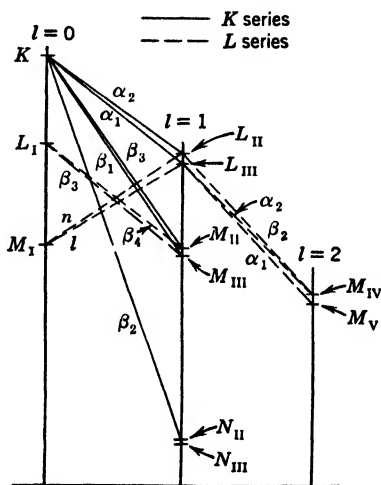


FIG. 18.4. X-ray energy-level diagram, showing allowed transitions giving rise to the principal lines in the K series and some of the lines in the L series.

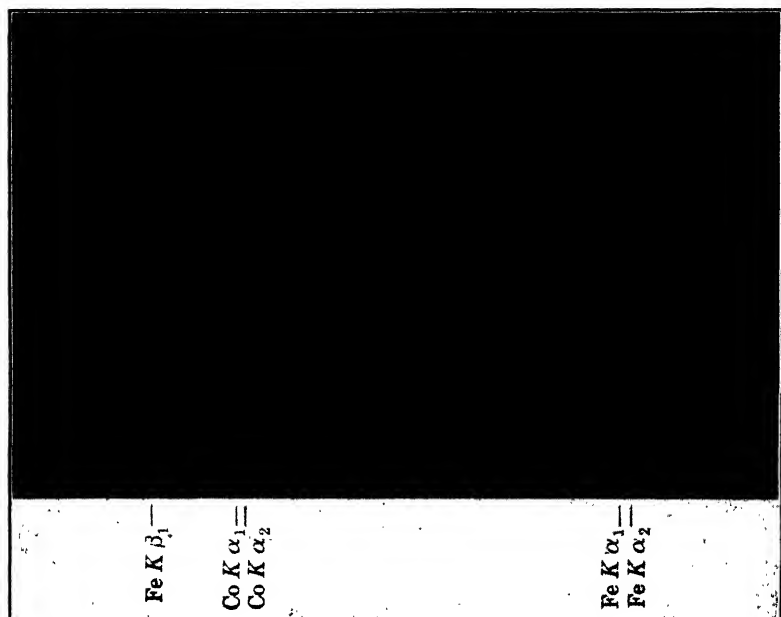


FIG. 18.5. Iron K -series lines and the cobalt K alpha doublet. Note the characteristic appearance of the K alpha doublets in this and the following spectrogram.

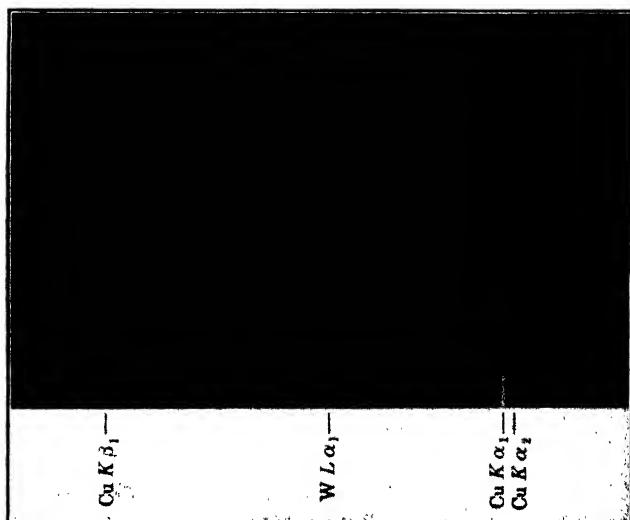


FIG. 18.6. K -series lines from a copper target. The tungsten L alpha lines frequently appear in the radiation from a Coolidge tube because of evaporation and sputtering of tungsten from the filament.

18.3 Non-diagram Lines

In general, experimental observations are in excellent agreement with the above energy-level scheme. However, now that the theory is well established, the occasional discrepancies are of

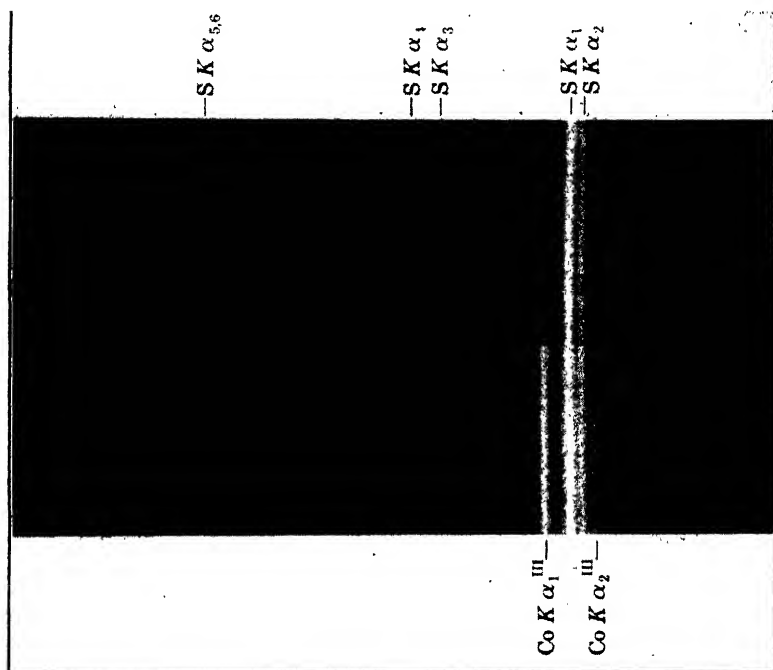


FIG. 18.7. Sulphur K alpha lines. The K alpha lines with subscripts 3, 4, 5, 6 are satellites radiated by atoms which have lost an L as well as a K electron. The cobalt K alpha lines in the third order may be used as convenient standards in measuring the wave lengths of the sulphur K alpha lines.

greater interest than its successes. There often appear spectrum lines which cannot be fitted into this energy-level diagram, particularly if one is studying the radiation from the light elements. These lines appear as weak lines which are usually close to a diagram line. Hence they are called **satellites** or **non-diagram lines**. They are of two kinds: those which appear on the long wave length side, and those which appear on the short wave length side. The former appear in the spectra of the so-called transition group elements such as Va , Cr , Mn , Fe , Co , Ni , or in the Pt , Pd

group, or in the rare earth group. In the K series of the iron group, the alpha lines are merely broadened toward longer wave lengths, the resolving power in Fig. 18.5 being too small to show this effect. The beta lines, however, have a weak satellite line which is fairly broad, and not very well resolved until the atomic number is lower than that of copper.

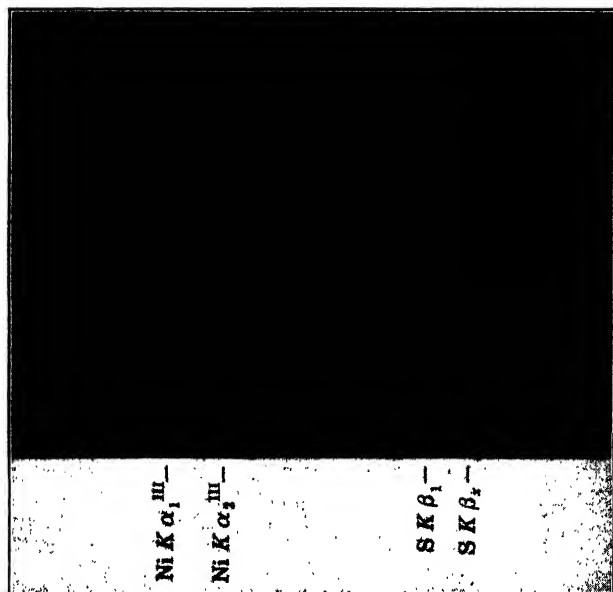


FIG. 18.8. Sulphur K beta lines. The line $K \beta_z$ is a short-wave satellite line. The nickel K alpha lines in the third order provide standard wave lengths. The K beta lines of light elements are sensitive to changes in chemical combination.

The satellites of short wave length are particularly intense and well separated in the elements of low atomic number, where the characteristic rays are so soft that a vacuum spectrograph must be used for their observation. Figures 18.7 and 18.8 are, respectively, spectrograms of the K alpha and of the K beta lines of sulphur taken with such an instrument. Of the seven lines of sulphur in these two figures (the original plates show nine), only three fit the energy-level diagram for sulphur shown in Fig. 18.9. Incidentally, the doublet structure of the β_1 line in the spectra of

elements of such low atomic number cannot be resolved because of the broadness of its component.

It will be noted on inspecting the diagram for sulphur and the configuration in Table 18.1 that the M electrons, which are involved in the emission of the beta lines, are in the valence shell of that element. It would naturally be expected that the wave

length, and perhaps the structure, of at least the beta line would be dependent on chemical combination.

During the first decade of x-ray spectroscopy it was thought that x-ray spectra were atomic in nature and were not influenced by chemical combination. The first positive evidence for chemical effects was obtained

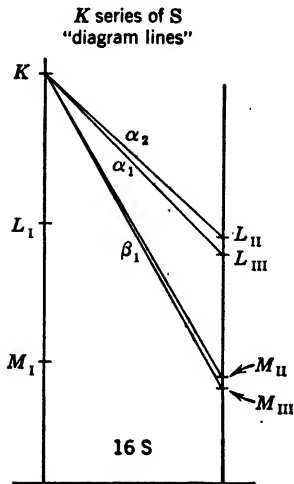


FIG. 18.9. X-ray energy-level diagram, showing the origin of the K -series "diagram lines" in the spectrum of sulphur.

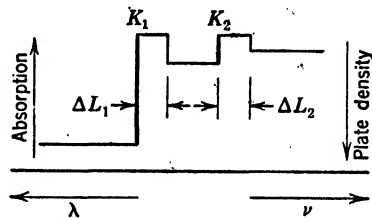


FIG. 18.10. Structure in the x-ray absorption spectrum of sulphur near the K edge.

by Bergengren in observing the absorption edges of phosphorus. This work has been extended by Lindh, Stelling, Lundquist, and others. Absorption edges give more clear-cut, unambiguous results than emission lines. For sulphur the general nature of the K absorption edge is as shown in Fig. 18.10. The wave lengths of the two edges K_1 and K_2 and the widths of the two absorption lines are as follows for sulphur:

	K_1	K_2	ΔL_1	ΔL_2	ΔK_1
Crystalline S:	5008.8 XU	4994.1	8.1	7.9
Divalent S:	5009.3
Tetravalent S:	4996.0	4988.1	4.3	4.2	6.3 volts
Hexavalent S:	4987.9	5.2	...	10.3 volts

The last column gives the displacement of the K_1 absorption edge in electron volts from that of crystalline sulphur.

Chemical effects in the K emission spectra of P, S, and Cl were first observed by Lindh and Lundquist in 1924. Further experiments were made by Backlin, Wetterblad, Valasek, and others. The K alpha lines of sulphur are found displaced by about 3 XU toward shorter wave lengths in the sulphates of Ca, Mg, Sr, and Ba. Backlin was able also to show a simultaneous displacement of the short-wave satellites or "spark" lines. The experiments are complicated by chemical changes occurring under electron bombardment, making it preferable to use a fluorescence method of excitation. Even this is not entirely free from objection. The principal difficulty lies in obtaining sufficient intensity without destroying the compound. The irradiated compound must be very close to the focal spot and must be replaced frequently.

18.4 General Radiation

X-rays are also emitted by electrons from the cathode when they are decelerated on striking the target. The initial energy of the impinging electrons is equal to

$$W_i = \frac{eV}{300} \text{ ergs}$$

V being the difference in potential between the cathode and the target in volts, while e is the electronic charge in esu. The energy just after the first collision, W_a , is that of some unstable, hyperbolic orbit. Since the latter is not quantized, the energy W_a may have any one of a continuous range of values from zero to nearly W_i . By the Bohr formula, the frequency radiated is

$$\nu = \frac{1}{h} (W_i - W_a)$$

and, because of the indefiniteness of W_a , it may have any value from zero to a maximum of

$$\nu_m = \frac{W_i}{h}$$

The spectrum is accordingly continuous with a high-frequency or **short-wave limit**. The intensity distribution curve with respect

to wave length has a maximum value at about $3\frac{1}{2}\lambda_{\min}$. Radiation due to the process described above is called **general radiation**. Since its spectrum is somewhat similar to that of white light from an incandescent solid, it is sometimes called "white x-radiation." Although the spectrum does not depend qualitatively on the nature of the target, its intensity increases with the atomic number. For this reason, a heavy element such as tungsten—which also has a high melting point—is often used as an efficient emitter of general radiation. The total, integrated intensity from an x-ray tube is largely general radiation. Although the characteristic lines are intense enough to be very prominent in x-ray spectra, they are extremely narrow in wave-length spread and actually contain relatively little energy.

PROBLEMS

1. Compute the approximate wave length of the L alpha line in the x-ray spectrum of tungsten, $Z = 74$, assuming a screening constant of six.
2. Verify the formula for the x-ray levels of heavy atoms:

$$\frac{\tilde{\nu}}{R} = \frac{(Z - s)^2}{n^2} - \alpha^2 \frac{(Z - d)^4}{n^3} \left(\frac{3}{4n} - \frac{1}{j + \frac{1}{2}} \right)$$

Derive from this (a) Moseley's law and (b) the spin doublet law. Spin doublets have the same quantum numbers n and l , but differ in j .

3. (a) What is the shortest wave length of x-rays emitted by a tube operating at 400 kv? (b) Explain the basis of the formula used, and derive its numerical constant from the values of h , e , and c .

4. The mass absorption coefficient of copper is 25 for a wave length of 0.5 \AA . What fraction of the incident radiation of 0.25 \AA wave length will be transmitted by a sheet of copper 0.5 mm thick? Density of copper = 9.0 gm/cm^3 .

5. X-rays having a wave length of 208.6 XU strike a block of copper which has energy levels at 661.6 and 81.0 in Rydberg units. What kinetic energies are to be expected in the secondary beta rays (photoelectrons) from these levels? Rydberg's constant = $109,737 \text{ cm}^{-1}$.

Chapter 19. Magneto- and Electro-optics

19.1 Faraday Effect

In searching for effects of magnetic and electric fields on light, Faraday discovered in 1845 the rotation of the plane of polarization of light passing through substances placed in a magnetic field with the light propagated in the direction of the lines of force. The effect was greatest in a very heavy flint glass that Faraday had made. The rotation is usually proportional to the field strength H and to the thickness of material traversed; but in films of ferromagnetic metals it is proportional to I , the "magnetization." The constant of proportionality per unit thickness and field strength is called the **Verdet constant** and may have either

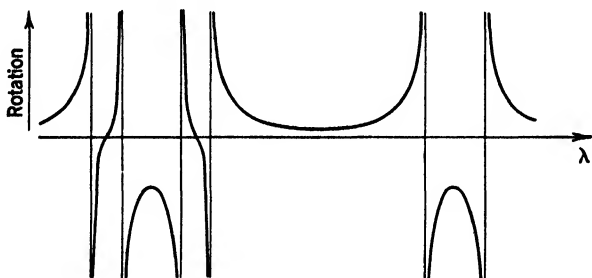


FIG. 19.1. Rotation of the plane of polarization of light passing through sodium vapor in the direction of a superposed magnetic field for wave lengths in the neighborhood of the two D lines. The lines show the longitudinal Zeeman splitting.

a negative or a positive value. The plane polarized light is sent through the substance parallel to the lines of force by placing the sample inside a solenoid or by using an electromagnet with holes bored longitudinally through the pole pieces. Macaluso and Corbino discovered that the rotation is quite large in sodium vapor near the D lines. The dispersion of rotation ρ was thoroughly investigated by R. W. Wood, who observed four complete revolutions near the D lines and verified the dispersion formula

$$\rho = \frac{a\lambda^2}{(\lambda^2 - \lambda_1^2)^2} + \frac{b\lambda^2}{(\lambda^2 - \lambda_2^2)^2}$$

where λ_1 and λ_2 are the wave lengths of the two D lines. Since each D line splits into several components in a magnetic field, Art. 19.4, the effect within the pattern of lines is really not so simple as the above formula indicates. Observing along the lines of force, one finds that λ 5890 (D_2) splits into four components, while λ 5896 (D_1) splits into two. The dispersion near and between the components was investigated by Zeeman in 1902 and found to be as predicted by Voigt in 1899. Qualitatively, the dispersion of rotation is as shown in Fig. 19.1.

19.2 Kerr Magnetic Effect

In 1875, Kerr observed that plane polarized light reflected from polished, magnetized pole pieces of an electromagnet has its state of polarization altered. It has been shown that plane polarized light reflected even off non-magnetized metals becomes elliptically polarized unless it vibrates in the plane of incidence or at right angles thereto. When the metal is magnetized, then even in these two cases the reflected light becomes elliptically polarized. This phenomenon is called the **magnetic Kerr effect**. The ellipse produced is usually very narrow and has its major axis inclined to the original direction of vibration. Hence the effect may be superficially regarded as a rotation of the plane of polarization.

19.3 Magnetic Double Refraction

When light passes through the material at right angles to the magnetic lines of force, there is a double refraction which depends in amount on the proximity to an absorption band. This effect was also predicted by Voigt in 1899. In sodium vapor the birefringence varies with wave length in the general manner shown in Fig. 19.2. When viewed transversely to the field, the D_2 line splits into six components, while the D_1 line splits into four, as shown.

The magnetic double refraction by liquids is not so pronounced as in vapors near their absorption lines, but was demonstrated by Cotton and Mouton in 1907. The effect is most easily observed in nitrobenzol and in carbon disulphide. Although the birefringence is small, it is about 1000 times larger than one would

expect from the associated Faraday effect, according to Voigt's earlier theory. However, in 1912 Voigt extended his theory to account for the discrepancy by making use of Langevin's theory of the orientation of molecules in a magnetic field. The lining up

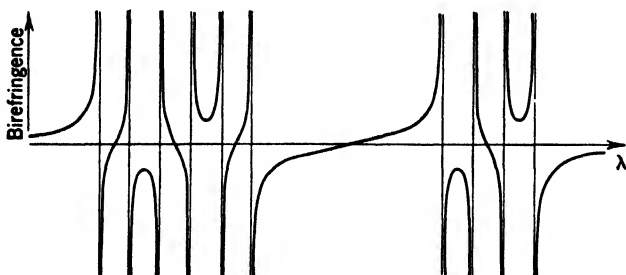


FIG. 19.2. Birefringence of sodium vapor when light is transmitted perpendicularly to the magnetic lines of force under the same conditions as in Fig. 19.1.

with the field is interfered with by thermal agitation, so that the Cotton-Mouton effect depends strongly on temperature, whereas the Faraday effect does not. Thus this effect is essentially different from that first predicted by Voigt.

19.4 Zeeman Effect

An effect by a magnetic field on the emission of light was first looked for by Faraday in 1862 but without success. Since the dispersion theory of Lorentz suggested the existence of a change in frequency of light emitted by electrons in a magnetic field, Zeeman tried again to observe such an effect. His first attempts were not successful, owing to the low field strengths used and the insufficient resolving power of his spectroscopic apparatus. However, in 1897 a broadening of the sodium D lines was observed. On Lorentz's suggestion Zeeman tried and verified the existence of polarization at the edges of the broadened lines. After improvements in technique, the lines were resolved into separated polarized components. When the light travelled at right angles to the field, Zeeman observed three plane polarized components; the central one p of unaltered frequency having its electric vector vibrating in the direction of H , and two outer s components which were polarized at right angles to p . When the light travelled along the lines of force through holes bored in the pole pieces of the

electromagnet, only the outer components were observed, and they were now circularly polarized in opposite senses. The relation between the sense of rotation of the electric vector and the direction of the field is shown in Fig. 19.3, where the magnetic field is directed away from the reader in Fig. 19.3(b) and toward the right in Fig. 19.3(a).

Lorentz was able to explain all the details of this normal Zeeman effect by his classical electron theory. The vibrations of the electrons in the magnetic field are resolved into components in the

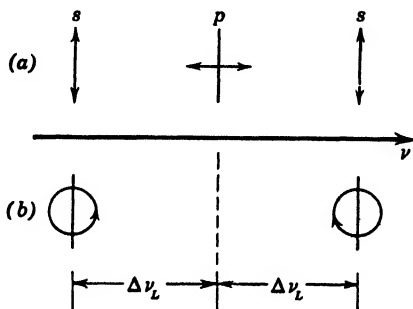


FIG. 19.3. Locations and polarizations of the components in the normal Zeeman effect. (a) When the light from the source travels perpendicularly to the magnetic field, three components are seen. (b) In the direction of the lines of force only the s components are seen and they are circularly polarized in opposite senses.

direction of the field and perpendicular to it. The frequencies of vibrations in the direction of the field are unchanged, since the field exerts no force on motions in this direction. The vibrations at right angles to the field may be resolved into two circular motions rotating in opposite senses. The frequencies of such motions are altered. Light emitted in the direction of the field is circularly polarized; while at right angles to it, the circular vibrations, seen projected in their own plane, emit light vibrating at right angles to the lines of force. The components of vibrations along the lines of force, which have their frequency unchanged, do not emit light in the longitudinal direction, but give the central line in the transverse effect. Even the correct magnitude of the shift was derived by Lorentz as follows. The equation of motion of an electron of mass m moving with a velocity v_0 in a circular path of radius r in the absence of a magnetic field is

$$\frac{mv_0^2}{r} = m\omega_0^2 r = br$$

whence

$$\omega_0 = 2\pi\nu_0 = \sqrt{\frac{b}{m}}$$

the constant b being a **force constant** and ν_0 being the frequency in the absence of the magnetic field. Assuming for the moment a positive charge rotating clockwise in a magnetic field directed toward the reader, the restoring force br is augmented by the amount Hev , where H and e are both in electromagnetic units. Thus the equation of motion becomes

$$m\omega_1^2 r = br + Hcv$$

and

$$\omega_1^2 = \omega_0^2 + H \frac{e}{m} \omega_1$$

or

$$\omega_1^2 - \omega_0^2 = H \frac{e}{m} \omega_1$$

The shift in frequency, which is relatively small, is thus seen to be

$$\nu_1 - \nu_0 = \frac{H}{4\pi} \frac{e}{m}$$

or the frequency itself is

$$\nu_1 = \nu_0 + \frac{He}{4\pi m}$$

For a counterclockwise rotation, similarly,

$$\nu_2 = \nu_0 - \frac{He}{4\pi m}$$

Since exactly the opposite shifts are observed, the latter or *right-hand* circularly polarized component, Art. 12.11, having the *higher* frequency, the vibrating charges must be of negative sign. Moreover, the value of e/m computed from the magnitude of the shift is about 10^{-7} emu/gm, showing that the radiating charges are electrons.

Soon after this satisfying correspondence between experiment and theory, it was found that the Zeeman effect is usually not so simple. For example, the sodium λ 5890 line actually separates into six components, while the λ 5896 line separates into four. The mercury green line at λ 5461 separates into nine components, the λ 4358 line into six, and the λ 4047 line into three. Although in the last case the pattern appears to be the same as in the normal effect, the frequency shifts are twice as large as predicted by the Lorentz theory. However, in most cases, the frequency differences can be expressed as integral multiples of fractional parts of the Lorentz splitting. This is the **anomalous Zeeman effect**.

In very strong fields the anomalous Zeeman effect approaches the normal splitting originally observed, a phenomenon known as the **Paschen-Back effect**. Except for light atoms, however, which have small multiplet separations, the field necessary becomes impossibly great. The rules for determining the weak field effect, or anomalous pattern, are readily stated, but the development of the formulas will not be given here. In the first place, one forms the sequence of magnetic quantum numbers M from the quantum number J (subscript of the term symbol) by the rule

$$M = -J, -J + 1, -J + 2, \dots, +J$$

This is applied to each of the two levels, the theory being that the angular momentum vector precesses about the magnetic field in the various quantized states M . Secondly, one computes the Landé splitting factor g by the formula

$$g = 1 + \frac{J(J + 1) + S(S + 1) - L(L + 1)}{2J(J + 1)}$$

The spin quantum number S in the formula is $\frac{1}{2}$ for single-electron spectra and either 0 (singlets) or 1 (triplets) for two-electron spectra. In general, S may be obtained from the multiplicity superscript, which is $2S + 1$. The orbital quantum number L is obtained from the letter symbol of the term, as noted in Art. 17.2. Each new level in the magnetic field is then found by applying a displacement of Mg Lorentz units to the original term, a **Lorentz unit** being

$$\frac{ehH}{4\pi mc} \text{ ergs} \quad \text{or} \quad \frac{eH}{4\pi mc^2} \text{ cm}^{-1}$$

The allowed transitions for which $\Delta M = 0, \pm 1$ give the anomalous or "weak field" pattern. As an illustration, the process is applied below to the mercury triplet:

$$7^3S_1 - 6^3P_{0,1,2}$$

$$\text{For } ^3S_1: \quad M = -1, 0, +1 \quad g = 2$$

$$Mg = -2, 0, +2$$

$$\text{For } ^3P_0: \quad M = 0 \quad g = \frac{3}{2}$$

$$Mg = 0$$

$$\text{For } ^3P_1: \quad M = -1, 0, +1 \quad g = \frac{3}{2}$$

$$Mg = -\frac{3}{2}, 0, +\frac{3}{2}$$

$$\text{For } ^3P_2: \quad M = -2, -1, 0, +1, +2 \quad g = \frac{3}{2}$$

$$Mg = -3, -\frac{3}{2}, 0, +\frac{3}{2}, +3$$

The resulting levels and allowed transitions are shown in Fig. 19.4, with the anomalous patterns indicated to a scale of Lorentz

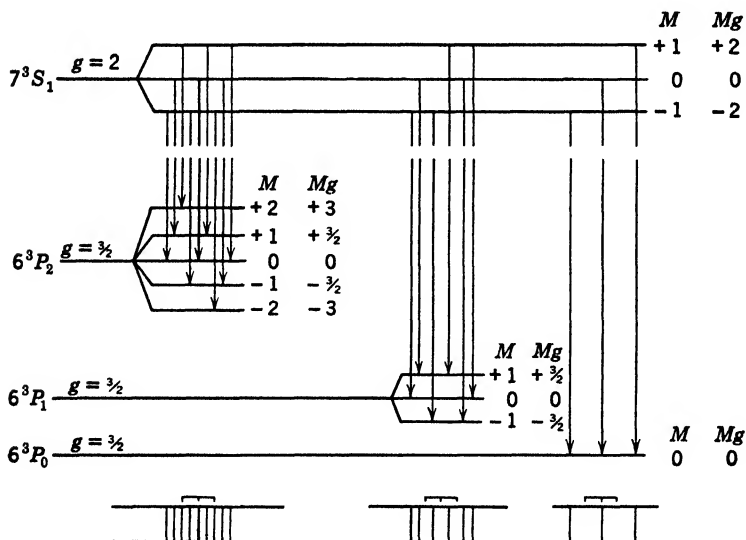


FIG. 19.4. Transverse anomalous Zeeman effect for the mercury triplet $\lambda 5461$, $\lambda 4358$, and $\lambda 4047 \text{ \AA}$. The scale above the groups of lines below the arrows indicates Lorentz units of separation.

units. The lines for which $\Delta M = \pm 1$ are circularly polarized in opposite senses when viewed in the direction of H . Viewed transversely to H , these same lines are plane polarized with their electric vector at right angles to H , while the lines for which $\Delta M = 0$ vibrate with their electric vector in the direction of H . These are the s and p components, respectively. The Zeeman pattern provides the spectroscopist with a means for determining or verifying the classification of spectrum lines. The normal pattern is observed only in the case of transitions between singlet states. It may, however, be observed with multiplets in the spectra of light atoms when the field is strong enough to produce a splitting which is large compared with the multiplet structure, as first observed by Paschen and Back.

19.5 Stark Effect

In 1913 Stark discovered that spectral lines radiated by a source in a strong electric field are also separated into components. There are both p and s polarized components, and the frequency shifts are at first proportional to the square of the field strength. A source for observing this effect is shown in Fig. 19.5. The positive ions or "canal rays" pass through holes in the cathode C into the

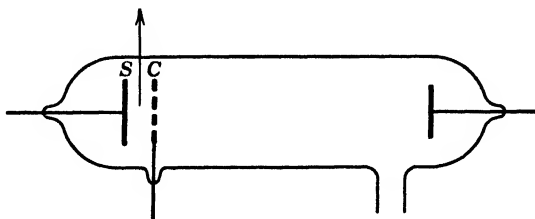


FIG. 19.5. A discharge tube for observing the Stark effect.

region CS , in which there is a strong electric field. The light from this region is analyzed spectroscopically. The Stark effect is completely explained by quantum mechanics and has no classical analog. The shifts are often unsymmetrical, so that lines appear displaced and broadened if the field is so small that the components cannot be resolved.

19.6 Kerr Electro-optic Effect

Transparent media become doubly refracting when subjected to a strong electric field, an effect discovered by Kerr in 1875. The

optical properties of the media are similar to those of uniaxial crystals, the optic axis, in this case, being in the direction of the field. The double refraction is said to be positive or negative, according to its resemblance to that of a positive or negative uniaxial crystal, Art. 12.9. When solid dielectrics are used, it is important to differentiate between a real electric effect and a double refraction due to stresses. For this reason, the most clear-cut experiments are with liquids in which there can be no stresses. Moreover, a rise in temperature can only change the index of refraction

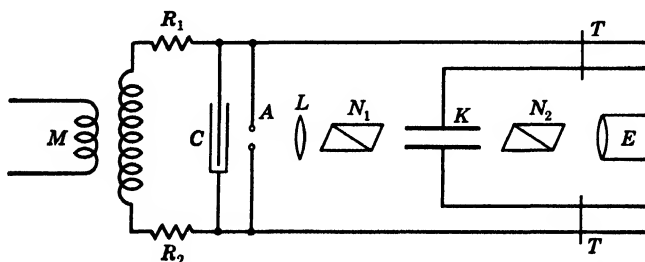


FIG. 19.6. Electro-optical shutter for transmitting flashes of short duration from the spark A.

but cannot lead to double refraction. The phase difference Δ is proportional to the square of the electric field strength E and the length of path l . Thus

$$\Delta = k l E^2$$

the birefringence being

$$n_p - n_s = j \lambda E^2$$

where j is the **Kerr constant**. For carbon disulphide and at a wave length of 6800 \AA , the constant j has the relatively large value

$$j = 3.70 \times 10^{-7}$$

in cgs units, the field E being in statvolts per centimeter. In this case the effect is **positive**, the ordinary ray, which vibrates at right angles to E , travelling faster than the extraordinary. Thus the corresponding index n_s is less than n_p .

Since the disappearance (or appearance) of double refraction follows a sudden drop (or rise) in the field by a lag of only 10^{-9}

sec, Kerr cells between crossed nicols have often been used as rapidly operating **electro-optical shutters**. Two arrangements by Beams¹ are shown in Figs. 19.6 and 19.7, the power supply across A being omitted in the second figure. N_1 and N_2 are two crossed nicols or Polaroid filters with their principal sections at 45° to the field E in the Kerr cell K . The shutter shown in Fig. 19.6 is open, because of the electric double refraction, until a spark passes at A . It then closes rapidly, the time interval between the passage of the spark and the closing of the shutter being varied by the adjustment of the trolleys T . The shutter in Fig. 19.7 has two Kerr cells K_1 and K_2 in which the fields are at right angles, these

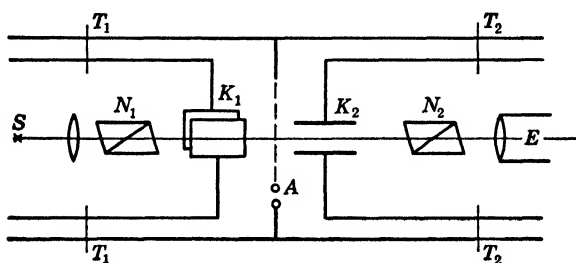


FIG. 19.7. Electro-optical shutter for transmitting flashes of short duration from any continuously emitting source S .

being placed between two polarizers N_1 and N_2 , which are crossed and set at 45° with the fields. The double refractions in the two cells balance when the field is at any steady value or is changing slowly. When, however, a spark jumps across A , the shutter is opened quickly as the double refraction in K_1 disappears, and is then closed rapidly when the double refraction in K_2 vanishes a little later, assuming that the trolleys T_2 are farther out than T_1 . A pulse of light as short as 1.5 meters can thus be transmitted, the time of the flash being about 0.5×10^{-8} sec. A rifle bullet travelling at a speed of 500 meters per second moves only $1/400$ mm in this time interval.

In piezoelectric crystals the electric double refraction is proportional to the first power of the field strength and is ten or more times as large as the ordinary Kerr effect. A good summary of this topic may be found in Cady's *Piezoelectricity*, pages 721–727.

¹ BEAMS, *Journal of the Optical Society of America and Review of Scientific Instruments*, **13**, 597 (1926).

19.7 Photoelectric Effects

When light of sufficiently high frequency strikes any material, electrons are ejected from it. This is the electron **photoemission** effect discovered by Hertz in 1887. In accordance with the quantum theory, the energy transferred in each act of absorption is $h\nu$. Of this, an amount w is required to liberate the emitted particle. The energy with which it leaves the surface is accordingly

$$E = h\nu - w$$

This equation is called **Einstein's photoelectric equation**. If w represents the binding energy of the electron, the observed energy E will in general be the *maximum* energy of the emitted particles, since many of the photoelectrons leave from atoms below the surface and travel at various angles with the surface, losing energy on the way out. The maximum energy E_m may be measured by means of the retarding potential V_m required to stop the electron emission. Then

$$eV_m = h\nu - w$$

This equation has been tested over a wide range of frequencies, and some very accurate determinations of h have been made by measuring the slope of the straight line relating V_m to ν .

The magnitude of w determines the photoelectric threshold, since it is obvious that no electrons may leave the surface if $h\nu < w$. Hence the **threshold wave length** is given by

$$\lambda_m = \frac{hc}{w}$$

The value of w , called the **work function**, is generally low for the free electrons in metals, particularly in the alkalis. The hydride of cesium has an unusually low value of w , and it is therefore widely used in photoelectric cells for detecting or measuring visible light.

The energy or velocity of the photoelectrons is entirely independent of the intensity of the light. The number of electrons emitted per second is, however, proportional to the intensity, making photoelectric cells very useful in the measurement of radiation. Unfortunately the constant of proportionality varies with the wave length, so that it is necessary to calibrate cells used in heterochromatic photometry.

Einstein pointed out that the energy transfer between radiation and electrons is most readily explained by a corpuscular theory of light. The corpuscles, called **photons**, carry an energy $h\nu$ and a momentum $h\nu/c$. The photon theory also explains the Compton scattering of x-rays very convincingly. However, the assumption of a quantized transfer of energy and momentum between electromagnetic waves and matter accomplishes the same result and is not so inconsistent with the phenomena of interference, diffraction, and polarization. It also enables one to retain the theory of propagation of light by which its speed in space c and in various media may be computed, and it leaves the theory of reflection and refraction intact.

The electrical conductivity of some semiconductors, such as selenium, decreases when they are illuminated. There is no simple relation between resistivity and intensity, but such cells may be very sensitive to radiation. They show enormous aftereffects, which exclude them from use in accurate comparisons of intensities.

Photovoltaic cells, such as the "photronic cell," consist of a thin layer of semiconductor between two metallic electrodes. The upper electrode is thin enough so that it transmits a considerable amount of light. Illumination of the cell causes a difference in potential between the two electrodes, arising from the easier passage of electrons from the semiconductor into the backing plate, or, under some conditions, from the semiconductor into the front electrode. Such a cell connected across a microammeter makes a conveniently compact unit for the measurement of illumination, being widely used as a foot-candle meter and as a photographic exposure meter. In the former application, the spectral sensitivity of the cell should approximate that of the human eye.

At x-ray frequencies one observes a photoelectric emission from the x-ray energy levels. One may measure the energy E of the ejected "secondary beta rays" by means of the radius of curvature of their path in a magnetic field. The binding energies of the various groups of electrons (K , L , M , etc.) are then obtainable from Einstein's equation

$$w = h\nu - E$$

When the quantum energy $h\nu$ exceeds twice the rest energy mc^2 of an electron, one observes the phenomenon of "pair pro-

duction," a positive and a negative electron being emitted simultaneously. This is sometimes referred to as "materialization of energy," since a pair of electrons, matter, appears and a quantum of radiant energy disappears in the process. The efficiency of the process increases with the energy. The threshold energy for this effect is about one million electron-volts (1 mev).

At higher energies, of over 2.19 mev, one may observe a photo-disintegration of the atomic nucleus with the emission of a neutron. This type of photoeffect has been most widely investigated with deuterium, in which it practically amounts to a fission of the nucleus. The threshold gives the binding energy of the proton and neutron.

The production of mesons by the interaction with nuclei of quanta having still higher energies is being investigated. Since the meson is a particle having a mass of the order of 200 times that of an electron, the threshold for such a process will be about 200 times the rest energy of the electron, or about 100 mev.

PROBLEMS

1. Using the formula for Landé's g factor, deduce the anomalous Zeeman pattern for the mercury 5770 Å spectrum line. Indicate polarizations of the components and their separation in Lorentz units.
2. Determine the field strength required to produce circular polarization by a Kerr cell filled with CS_2 if the path length is 10 cm.
3. Plane polarized light is transmitted forward and back through a plate of glass 2 cm thick in a magnetic field. The light travels along the magnetic lines of force, and the Verdet constant of the glass is 60.8×10^{-3} minute per gauss and cm. What field strength would be required to produce a rotation (Faraday effect) such that the light may be annulled on passing back through the original polarizer?

Chapter 20. Measurement of Color

20.1 Trichromatic Colorimetry

The wide use of color in photography and in industry makes it necessary to be able to specify color by numbers and to state the permissible tolerances in terms of deviations from these numbers. The measurements should be entirely objective, although the standards may have a subjective historical origin. A satisfactory process of color measurement is presented in Hardy's *Handbook of Colorimetry*, published by the Technology Press.

The specification of any color is most completely given by spectrophotometry, in which one measures the reflectance or transmittance of the sample as a function of the wave length. To obtain a second sample which will match the first under any kind of illumination, the entire spectrophotometric curve must be the same for both samples. For many purposes, however, a statement of three numbers will suffice, although these numbers will be different when different sources are used, even for the same sample, since it is known that colors of materials depend on the kind of illumination used. To reduce these possibilities, one generally selects one of three standard illuminants which have been adopted by the International Congress of Illumination (ICI). The so-called illuminant *A* is similar to light from a gas-filled tungsten lamp having a color temperature of 2848°K, *B* is an approximation to mean noon sunlight, and *C* is an approximation to average daylight. The International Congress has specified the spectral distribution of these three sources by means of graphs and tables. Illuminant *C* is used in most color work.

It is an experimental fact that colors may be matched by various proportions of three so-called **primary colors**. This fact is independent of any theory of color vision. For any set of three real primaries, however, it is occasionally necessary to use negative values of one or more of the primaries or, in other words, to add them to the color being matched. Contrary to popular impression, the choice of primaries is quite arbitrary, and one may readily

transform measurements made in terms of one set into values corresponding to any other set. It is only to reduce the likelihood of needing negative values that certain shades of red, green, and blue-violet are generally chosen as practical primaries. The ICI has specified three standard primaries which avoid negative values entirely. These are defined by tables giving the excitation values \bar{x} , \bar{y} , and \bar{z} of spectral wave lengths in terms of the three standard primaries. **Excitation values** are luminous intensities per unit amount of radiant flux. These values were derived by linear transformation from the results of experiments in matching spectral colors by three real primaries. The \bar{X} primary is a reddish purple of higher saturation than any real color, the \bar{Y} primary is a green of considerably greater saturation than λ 520 m μ , and \bar{Z} is a blue of greater saturation than λ 477 m μ . These conclusions are obtained by the procedure outlined in the next article. For the present, the primaries may be considered to be defined by the excitation values specified by the ICI and tabulated in Hardy's *Colorimetry*.

To make a trichromatic analysis of a color, one first determines the spectral reflectance R_λ of the sample by means of a spectrophotometer. If the sample is a transparency, one substitutes the transmittance T_λ for R_λ . The **tristimulus values** of the color, using a source whose intensity distribution is given by J_λ , are defined by the equations

$$X = \lim_{\Delta\lambda \rightarrow 0} \sum J_\lambda R_\lambda \bar{x}_\lambda \Delta\lambda$$

$$Y = \lim_{\Delta\lambda \rightarrow 0} \sum J_\lambda R_\lambda \bar{y}_\lambda \Delta\lambda$$

$$Z = \lim_{\Delta\lambda \rightarrow 0} \sum J_\lambda R_\lambda \bar{z}_\lambda \Delta\lambda$$

These summations may be replaced by integrals, but in practice one finds it more convenient to use small values of $\Delta\lambda$ and to carry out the indicated summations on an adding machine. Since the values of R_λ may be obtained by means of a photoelectric recording spectrophotometer and J_λ and $\bar{x}\bar{y}\bar{z}$ are adopted standard functions, it is evident that the tristimulus values are derived by an objective process in which the color vision or even the possible color blindness of the operator plays no part.

20.2 Chromaticity

The quality of a color or its **chromaticity** is defined by the ratios

$$x = \frac{X}{X + Y + Z}, \quad y = \frac{Y}{X + Y + Z}, \quad z = \frac{Z}{X + Y + Z}$$

Since $x + y + z = 1$, only two of these are independent, so that x and y are sufficient to designate chromaticity. Moreover, since the adopted \bar{g}_λ function has the same values as the visibility

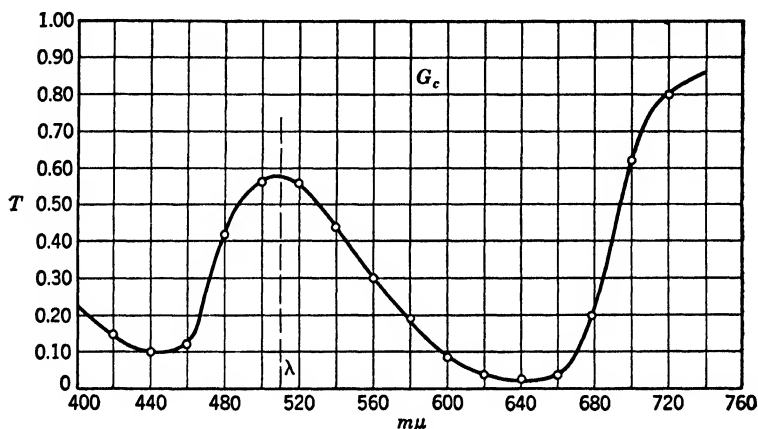


FIG. 20.1. Spectrophotometric transmissivity curve for a sample of green Cellophane.

function, the tristimulus value Y is proportional to the **luminosity** as given by the overall reflectance or overall transmittance of the sample, depending on the nature of the material. Thus x , y , and Y are three numbers which suffice to designate the color of any sample. From these values one may find the **dominant wave length** or hue, the **saturation**, and the **luminosity** of the sample, as will be shown below.

As a concrete example, the spectrophotometric transmission curve for a sample of green Cellophane is given in Fig. 20.1. The tristimulus values found by summing the products $J_\lambda T_\lambda \bar{x}_\lambda$, etc., at 10-mμ intervals are

$$X = 157.96, \quad Y = 324.88, \quad Z = 243.17$$

Hardy explains how these values may be found more easily by summing values of T_λ at selected ordinates which he tabulates. Since the sum Y for a perfectly transparent sample, $T = 1$ for all wave lengths, is 1065, the relative luminosity of the green is

$$\frac{324.88}{1065} = 0.305 \quad \text{or} \quad 30.5 \text{ per cent}$$

The chromaticities are found to be $x = 0.2174$ and $y = 0.4480$. Figure 20.2 shows a color chart or color triangle, in which this point

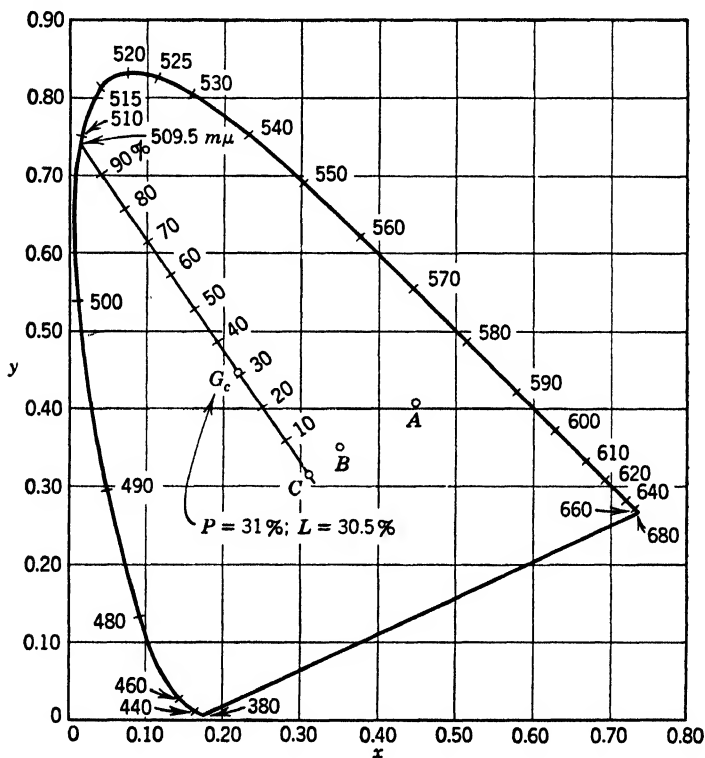


FIG. 20.2. Chromaticity plot after Hardy. The point G_c indicates the chromaticity of the sample of green Cellophane whose transmissivity is graphed in Fig. 20.1.

is plotted at G_c . The curved line is the locus of chromaticities of spectral colors, while C is the point representing illuminant C , which was used in the computations which employed the values of

J_λ of this source. A straight line through G_c from C intersects the locus of spectral hues at λ 509.5 m μ . This is the dominant wave length. Since G_c is 31 per cent of the way from C to the spectral locus, the purity of this green is 31 per cent. The luminosity is given as before by Y , as compared with a perfectly transparent sample, or $L = 30.5$ per cent. Thus one may specify the color of this sample by giving either the three numbers $x = 0.2174$, $y = 0.4480$, $L = 30.5$ per cent, or λ_{dom} 509.5 m μ , $P = 31$ per cent, and $L = 30.5$ per cent. When necessary, one may specify acceptable tolerances in these numbers, making it possible to designate any desired color with any degree of precision.

The nature of the ICI primaries for which x , y , z are respectively 100, 010, and 001 may be analyzed by the above procedure, which enables one to determine their dominant wave length and saturation. Since the primaries lie outside the spectral locus, their saturation is found to be greater than 100 per cent. Consequently they are not real colors and cannot be duplicated by any dyes.

In practical matching with three colors, one must use three real primaries, which are represented by three points somewhere within the spectral locus. Positive amounts of these in various proportions enable one to match any color within a triangle formed by connecting the three points. The greatest variety of colors can be so matched if the triangle has as large an area as possible. It is for this reason that one uses some shade of red, green, and blue-violet. The type of color mixture considered here is not directly comparable to that obtained by the mixing of dyes, referring rather to the addition of colored lights.

20.3 Color Mixture

If unit amounts of three primary colors have the tristimulus values $X_r Y_r Z_r$, $X_g Y_g Z_g$, and $X_b Y_b Z_b$, the tristimulus value of an **additive mixture** of r units of the first, g units of the second, and b units of the third is given by

$$X' = rX_r + gX_g + bX_b$$

$$Y' = rY_r + gY_g + bY_b$$

$$Z' = rZ_r + gZ_g + bZ_b$$

From these one may obtain the chromaticity of the mixture by the usual method. As mentioned before, all chromaticities within a

triangle bounded by lines joining the chromaticities of the three primaries may be duplicated by varying the amounts r , g , b of the primaries. This principle is employed in color photography and other methods of color reproduction.

In the case of additive processes, the application is quite straightforward, as outlined above. In **subtractive processes**, however, the light passes through the three colored images in succession. In this case, one must think of an increase in concentration of each dye as causing a *reduction* in the amount of its corresponding primary. Each of the dyes used in such subtractive combinations must have a color which is complementary to the primary color which it controls. Thus a cyan dye controls red, a magenta dye controls green, and a yellow dye controls blue. In a sense, the dye colors may be called the primaries for subtractive processes, but it is preferable to consider the colors which they absorb. It is evident that dyes do not add any light and therefore color; they merely remove a fraction of what is there already, the fraction depending on their concentration. To be able to duplicate the greatest variety of colors by mixtures of such dyes, one must select three pigment colors complementary to the three additive primaries which serve as corners of the largest possible triangle on the chromaticity diagram. By using optical densities of unit dye concentrations instead of the tristimulus values X , Y , Z , one may write equations for the subtractive process which are exactly like those at the foot of page 329. Further discussion of the theory of three-color reproduction may be found in an article by Hardy and Wurzburg in the *Journal of the Optical Society of America*, **27**, 227-240 (1937) and in Miller's *Principles of Photographic Reproduction*.

PROBLEMS

1. Compute the tristimulus values X , Y , Z and from these the trichromatic coefficients x , y , z for the material whose spectrophotometric curve is given in Fig. 20.1, using illuminant A. Reference: HARDY, *Handbook of Colorimetry*.
2. For the same illuminant and material as in problem 1, determine the dominant wave length, purity, and luminosity.
3. Determine the relative amounts of the three real primary colors 0.60, 0.30; 0.10, 0.70; 0.10, 0.20 needed to match the chromaticity of the material whose spectrophotometric curve is given in Fig. 20.1, using illuminant C.
4. Prove that the additive mixture of any two colors x_1 , y_1 and x_2 , y_2 lies on a straight line connecting the two points on the x , y diagram.

Chapter 21. The Scattering of Light

21.1 Tyndall Effect

The intensity of a beam of light passing through any medium is attenuated because of, first, a **true absorption** in which the radiant energy is converted into another form, and, second, a **scattering** of radiation in all directions. The scattering of light by small particles suspended in the atmosphere was first studied by Tyndall. He found that large particles scatter incident white light without altering its color, but that as the size of the particles becomes smaller than the wave length of light, the scattering increases as the wave length diminishes, causing the scattered light to acquire a blue color. He indicated that the blue color of the unclouded sky is due to this phenomenon. Lord Rayleigh, however, pointed out that even dust-free air scatters light because of molecular fluctuations in density and that the amount and quality of scattering due to this effect are sufficient to account for the blue color of the clear sky. When air contains fine particles of dust or smoke, the molecular scattering is augmented by the Tyndall effect. The light transmitted through long distances in the atmosphere, as at sunset, is then a brilliant red.

21.2 Rayleigh Scattering Formula

It has been proved, Art. 14.3, that the electric polarization P in a medium subjected to an oscillating field E of angular frequency ω is given by

$$\frac{P}{E} = \sum_h \frac{N_h e^2}{m_h (\omega_0 h^2 - j \omega_h' \omega - \omega^2)}$$

Hence, if one neglects absorption and assumes only one important natural frequency ν_0 , the polarization of a single particle is given by

$$P = \frac{e^2 E}{4\pi^2 m (\nu_0^2 - \nu^2)} = \frac{e^2 E}{m (\omega_0^2 - \omega^2)}$$

This polarization will vary with the frequency ν of E , leading to a scattering of light of this same frequency. The total rate of scat-

tered radiation in all directions may be proved to be related to P^2 in accordance with the formula

$$F_s = \frac{\omega^4 P^2}{3c^3}$$

the proof of which will be omitted here. On substituting the value of P and assuming the presence of N particles scattering non-coherently, so that one may add intensities, one finds that

$$F_s = \frac{Ne^4 E^2}{3m^2 c^3 \left(\frac{\omega_0^2}{\omega^2} - 1 \right)^2}$$

Since the average flux density of incident radiation is

$$I = \frac{c}{8\pi} E^2$$

the **scattering power**, or fraction of the incident radiant flux which is scattered, is given by

$$\frac{F_s}{I} = \frac{8\pi Ne^4}{3m^2 c^4 \left(\frac{\nu_0^2}{\nu^2} - 1 \right)^2}$$

If the natural frequency lies far enough in the ultraviolet, as is often the case in optical scattering, $\nu_0^2 \gg \nu^2$, then

$$\frac{F_s}{I} = \frac{8\pi Ne^4 \nu^4}{3m^2 c^4 \nu_0^4} = \frac{8\pi Ne^4}{3m^2 \nu_0^4 \lambda^4}$$

This is one form of Rayleigh's scattering formula. One notes that the scattering power varies as the inverse fourth power of the wave length, hence should be much more intense for blue than for red light.

On the other hand, if one is interested in x-ray scattering where $\nu^2 \gg \nu_0^2$, one finds that the scattering power approaches

$$\frac{F_s}{I} = \frac{8\pi Ne^4}{3m^2 c^4}$$

This is Thomson's classical formula for the scattering of x-rays.

Another form of Rayleigh's formula which gives, instead of the

integrated flux F_s , the intensity of the scattered radiation I_s , as a function of the scattering angle θ , is

$$I_s = I \frac{(n' - n)^2}{n^2} (1 + \cos^2 \theta) \frac{N\pi V^2}{r^2 \lambda^4}$$

where n' and n are the indices of refraction of the particles and of the medium respectively, V is the volume of each particle, λ is the wave length, and r is the distance from the scattering region to the observer. The factor $(1 + \cos^2 \theta)$ applies to unpolarized incident light. For components vibrating at right angles to the plane defined by the direction of the incident light and the direction of observation this factor is unity, while for components in this plane it is just $\cos^2 \theta$. Accordingly, the scattered light is always more or less polarized, with a maximum polarization for $\theta = 90^\circ$. This is easily verified in the case of light from the sky. Since the intensity varies inversely as the fourth power of the wave length, blue light will be scattered more intensely than red. As the particle size increases, the angular distribution and polarization change markedly. Particles which have strong absorption in the visible region of the spectrum will be found to deviate from the inverse fourth power law of ordinary Rayleigh scattering.¹

The blue color of the sky and the bluish haze which intervenes between an observer and distant hills or other objects are due to Rayleigh scattering, although often modified by the presence of large particles which scatter non-selectively. It is well known that a blue-absorbing or yellow filter may be used in photography to absorb much of this scattered light. A deep red or infrared transmitting filter cuts out the haze still more effectively. By the use of an infrared filter, black sky effects may be obtained in daylight, and mountains several hundred miles away may be photographed from balloons or airplanes.

It is possible to study scattering due to density fluctuations in a relatively small amount of any medium if precautions are taken to eliminate scattering by particles of foreign matter and reflection from the walls of the vessel or other surfaces. One might expect a greater scattering from liquids and solids than from gases because of the greater number of particles in a given volume. That this is not always the situation is due to the greater uniformity of density

¹ MIE, *Annalen der Physik*, **25**, 377 (1908).

in these media. In particular, transparent crystals such as quartz show very little scattering. Their uniform lattice structure leads to smaller density fluctuations than are found in a liquid or a gas, and thus they behave like a more homogeneous medium. The scattering by a vapor increases with the temperature up to the critical temperature, at which the density fluctuations are so great that they give rise to a phenomenon known as **opalescence**.

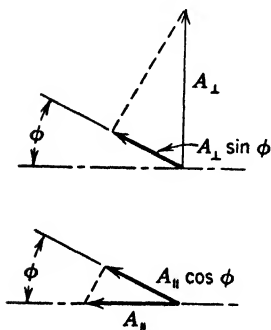


FIG. 21.1. Measurement of depolarization factor by means of the analyzer angle ϕ at which equal amounts of the two components A_{\perp} and A_{\parallel} are transmitted.

The light scattered at 90° by gases and liquids is usually only partially polarized. This is due to a molecular anisotropy, in consequence of which the induced dipole moments are not in the direction of the electric vector. Since the molecules are oriented in various directions, the scattered light becomes depolarized. The degree of depolarization of scattered light may be studied by a Wollaston double-image prism, oriented so that one of the images contains light vibrating at right angles to the plane defined by the incident ray and the direction of observation. If the light were completely polarized, only this single image would be observed, but actually one may see both. If one first rotates an analyzing nicol to extinguish

the first image (providing a "zero" setting), and then observes the change in angular setting ϕ required to match the intensities of the two images, he obtains a measure of the ratio of their original intensities. As can be seen from Fig. 21.1, the condition for a match is

$$I_{\perp} \sin^2 \phi = I_{\parallel} \cos^2 \phi$$

since intensity is proportional to the square of the amplitude. The ratio of the intensities of the light vibrating in the plane of scattering and at right angles to it is

$$\frac{I_{\parallel}}{I_{\perp}} = \tan^2 \phi \equiv r$$

The value of r , called the **depolarization factor**, would be zero for complete polarization and unity in the absence of polarization.

The depolarization factor for several liquids is given in the following table, which shows the wide range of values that may be encountered.²

DEPOLARIZATION FACTOR FOR LIQUIDS

Benzene	0.45	Chloroform	0.24
Chlorobenzene	0.50	Ether	0.09
Nitrobenzene	0.73	Carbon tetrachloride	0.045
Carbon disulphide	0.64	Methyl alcohol	0.062

21.3 The Raman Effect

In 1928 C. V. Raman discovered changes in frequency on scattering of monochromatic light. These frequency shifts are equal to differences between molecular energy states divided by Planck's constant. The most intense lines in the spectrum of the scattered light are generally those in which the frequency ν is diminished to ν' , called the **Stokes's lines**. The lines whose frequency ν'' after scattering is greater than that of the incident light are called **anti-Stokes's lines** and are generally much weaker. If M and M^* represent the ground state and an excited state of the molecule respectively, the scattering process in which a transition occurs from one of these states to the other must satisfy the principle of conservation of energy, which may be written in the form

$$M + h\nu = M^* + h\nu'; \quad \nu' < \nu$$

and

$$M^* + h\nu = M + h\nu''; \quad \nu'' > \nu$$

In order that either type of scattering may take place, at least one other molecular state must exist, which may combine with both M and M^* in accordance with the standard selection rules. On the other hand, direct transitions between M and M^* may be forbidden, making it possible to observe Raman shifts in frequency of $(M^* - M)/h$, even though these frequencies may not occur in the ordinary molecular absorption and emission spectra. The molecular states M and M^* are vibration-rotation states with the rotational fine structure usually unresolved, except in the case of gases. The rotational quantum number is observed to change by two units. No electronic Raman effect has ever been observed. Raman scattering differs from Tyndall or Rayleigh scattering in that the frequency of scattered monochromatic light is altered

² GANS, *Zeitschrift für Physik*, **17**, 353 (1923).

in the former, but not in the latter. It differs from fluorescence in that the incoming radiation is not actually absorbed by the molecule, and the frequency pattern of the Raman lines is independent of the frequency of the incident light. The intensity of Raman scattering, however, increases as the inverse fourth power of the wave length. Raman scattering is less intense than Rayleigh scattering, and for this reason it was not discovered earlier.

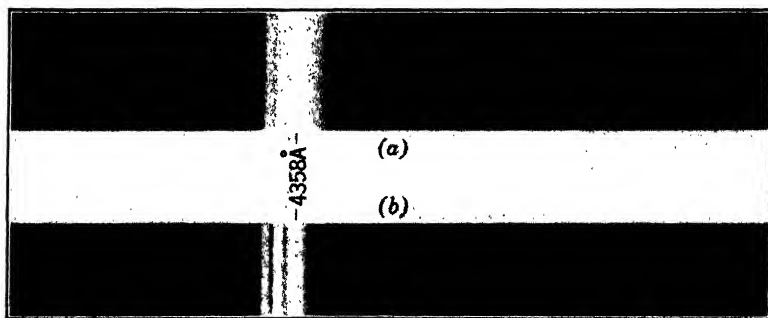


FIG. 21.2. Raman effect. The strongest line is the blue mercury line $\lambda 4358$, which is responsible for the Raman satellites. The neighboring mercury lines $\lambda 4347$ and $\lambda 4339$, seemingly strong because of overexposure, are actually too weak to show observable Raman lines in these spectrograms. The scattering was observed at 90° in (a) benzene, and (b) carbon tetrachloride. The lines of longer wave length, toward the right, are "Stokes" lines, while those of shorter wave length are "anti-Stokes" lines. (*Spectrograms by Sister Magna Werth, College of Saint Benedict.*)

In order to observe the Raman effect experimentally, it is customary to mount one or more intense mercury-arc lamps near a tube containing the scattering substance. The tube should have a horn-shaped, blackened end turned away from the spectrograph. The shape of this end causes a total reflection of backward scattered light and provides a dark background against which the effect is more easily observed. It is common practice to surround the scattering tube with a water jacket for cooling, or to use an air blast for this purpose. A filter of Noviol glass or a saturated solution of sodium nitrite is often used to absorb the ultraviolet, which may cause photochemical decomposition of the sample. Wood states that a thin sheet of pale Noviol combined with an aqueous solution of quinine sulphate acidulated with a few drops of sulphuric acid gives almost complete absorption of the ultraviolet,

including the strong mercury 4046 line, while the 4358 Å line is almost perfectly transmitted. A monochromatic filter arrangement such as this makes the Raman lines easier to identify, since one eliminates possible overlapping Raman patterns excited by neighboring strong lines. The spectrograms of Fig. 21.2 show the Raman patterns of benzene and of carbon tetrachloride produced principally by the scattering of light having a wave length of 4358 Å.

IV. EXPERIMENTAL OPTICS

Experiment 1. Cardinal Points of a Thick Lens

The thick lens in this experiment consists of a biconvex lens and a biconcave lens whose principal planes are engraved on their mounts. They are held coaxially by a cylinder, and their separation d , Fig. 1, may be set at any of the values indicated by the instructor. The combination illustrates a simple form of telephoto lens system. The object of this experiment is to locate the cardinal points experimentally and to compare the measured loca-

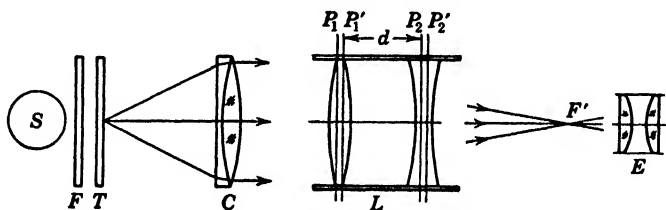


FIG. 1. Alignment of units on an optical bench for determining the location of the cardinal points of the lens combination L .

tions with the results of computation, using the formulas derived for combinations of optical systems in Art. 3.3.

The lens system is mounted in a nodal slide on an optical bench. The target T , Fig. 1, is illuminated by an incandescent lamp S , the light from which is made approximately monochromatic by the use of a red filter F in front of the target. With the aid of a telescope which has been focussed on a distant object, the distance between T and the lens C is adjusted so that the rays from any point of T are parallel and the target is virtually located at infinity. The image of the target created by the lens combination L will then be in its second principal focal plane. The lens L is rotated back and forth about a vertical axis and displaced along the nodal slide

until the image as viewed with the aid of the eyepiece E shows no transverse motion. The axis of rotation must then pass through the second nodal point N' , which is coincident with the second principal point P' since the lens system is in air. A consideration of Fig. 2 will show that the above statement is true, since any principal ray passing through N will emerge from N' in a direction parallel to its original course. Since the incident rays are parallel, there will always be some ray parallel to the optical bench and passing through N . The conjugate ray will always emerge along the axis from N' , as this point is fixed on the axis of rotation. Since the image must be on this principal ray, at a point where it is inter-

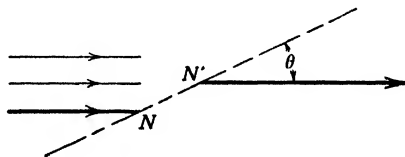


FIG. 2. Rotation of a lens system about an axis passing through the second nodal point N' does not cause a lateral displacement of a distant object.

sected by other rays, the image will not be displaced transversely by a rotation about N' . The location of the axis of rotation is indicated by a pointer on the nodal slide. Its distance p' from the last principal plane of the last component may be readily found by means of a divider and scale. The distance from N' to F' which is equal to the focal length f' is given directly by the distance between the stands holding the nodal slide and the cross-hairs in front of the eyepiece. By rotating the lens combination through 180° and repeating the above procedure, one may determine p and f . The values of f and f' should be equal within experimental error.

In order to compute f , f' , p , and p' , one must know not only d but also the focal lengths f_1 and f_2 of the component lenses. The focal length of the converging lens may be found by the same procedure as for the combination. The focal length of the diverging lens is determined from the conjugate distances s_2 and s_2' , Fig. 3, using the converging lens as an auxiliary lens so that the final image is real. With a little ingenuity, the distances s_2 and s_2' may be measured directly on the optical bench without making use of the focal length of the auxiliary lens. If the collimator lens C is removed, the distances are larger and may be measured more ac-

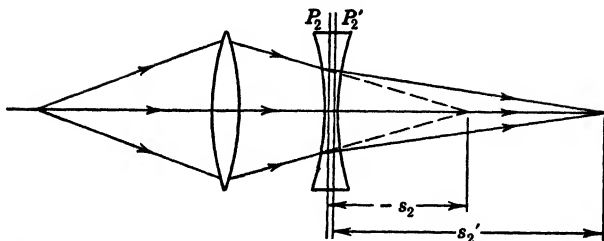


FIG. 3. Measurement of the conjugate focal distances s_2 (negative) and s_2' enables one to determine the focal length of a diverging lens.

curately. Since s_2 is a distance to a virtual object, its value is negative in the lens formula

$$\frac{1}{f_2} = \frac{1}{s_2} + \frac{1}{s_2'}$$

The values of p , p' , f , and f' should be measured and computed for at least two different separations d , for example, $d = 1$ cm and $d = 3$ cm.

Experiment 2. Aberrations of Lenses

The object of this experiment is to test an old-style photographic objective for the following aberrations: (a) spherical, (b) chromatic, (c) astigmatism, and (d) curvature of image. The lens is mounted in the nodal slide of an optical bench provided with a micrometer microscope M , Fig. 4, for measuring the displacements of the

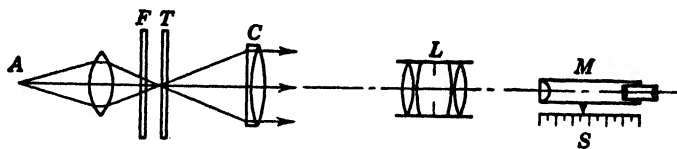


FIG. 4. Alignment of units on an optical bench when determining the aberrations of a lens system L .

image. In the figure the micrometer screw is represented by the scale S . The light from the mercury-arc source A is focussed on the target T through a filter F which isolates one of the principal

wave lengths, such as λ 5461. The collimating lens C and target T are first adjusted for parallel rays with the aid of an auxiliary telescope as in the preceding experiment. The use of the lens C enables one to place the target effectively at infinity while still having all parts on a bench of conventional length. If, however, the lens C is suspected of having appreciable aberrations of the kind for which tests are being made, it should be removed and the target placed at a distance of at least fifteen focal lengths from L . If this is inconvenient, the tests may be made at the maximum convenient distance, which should be stated, since the aberrations of a lens depend on the distance of the object.

The alignment of all units on the same optical axis should be first approximated by careful inspection. Rotation of the target T will serve as a check on the location of the intersection of its index lines at the center of rotation of the holder. By sliding the lens L along the bench and checking the transverse motion of the image, one may determine whether the object is on the principal axis of the lens L , or the secondary axis, if L is inclined. The lens L may then be displaced in the nodal slide until a rocking about the axis of rotation produces no lateral displacement of the image. This nodal point setting is not essential for the first parts of this experiment, but it may be conveniently made at this point, as it will be necessary in the test for astigmatism.

The longitudinal spherical aberration is measured by observing the difference between the focal lengths for paraxial rays and for rim rays. The iris diaphragm in the lens is closed to about one third of its full opening. Further closure may unduly increase the depth of focus, making the exact location of the image difficult to determine. The microscope is set on the image five times in the same direction to avoid backlash of the screw, the micrometer being read each time. Then the stop is opened wide, and the central four fifths or so of the area of the lens is blocked off by a circular diaphragm of black paper. The rim rays so obtained will be found to focus at a different place, for which the micrometer readings are found as before. The difference between the two sets of readings is the longitudinal spherical aberration, this being positive if the rim rays focus closer to the lens than the paraxial. A more complete test requires the isolation of successive zones.

To measure longitudinal chromatic aberration, one stops the lens down just enough to reduce spherical aberration to a scarcely

perceptible amount. The micrometer readings of the image location are made, using light of the following wave lengths in millimicrons: 436, 546, 578, and 650. The first three are obtained from the mercury arc by means of special filters called **mercury-vapor monochromats**. The 650 m μ red is obtained from an incandescent lamp, using a deep red filter. This light is not monochromatic but is approximately so, because of the sharp cut-off of the filter at its short wave limit and the steep drop in the visibility function of the eye at the red end of the spectrum not far from the cut-off. The change in focal length of the lens with color as com-

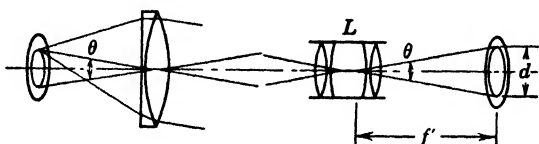


FIG. 5. Showing how an angular scale collimator enables one to determine the focal length of a lens system L by measuring the diameter of an image in its focal plane.

pared with its value for 546 m μ is plotted graphically to show the chromatic variation in focal length.

The focal length itself for 546 m μ is measured with the aid of an **angular scale collimator**. This consists of a target of concentric circles in the focal plane of a good telescope objective. The collimator is adjusted by sighting into it with a telescope which has been focussed on a distant object through an open window. The collimator lens is displaced until the image of the circles is in the same plane as the cross-hairs in the telescope. The collimator is then placed close to the lens whose focal length is to be measured. As shown in Fig. 5, the lens creates images of the circles in its own focal plane. By tracing principal rays one finds that these images subtend the same angle at the second nodal point N' of the lens as they do at the first nodal point of the collimator lens. The diameter of a ring subtending an angle θ is measured by means of the travelling microscope. If d represents the diameter of the ring, one readily finds that

$$f' = \frac{d}{2 \tan (\theta/2)} = kd$$

The value of k is a constant of the collimator and may be deter-

mined by measuring the angle θ by means of a spectrometer telescope or by any telescope mounted on an angular scale.

To test the lens for astigmatism one first aligns the units on the optical bench and displaces the lens in the nodal slide until the image shows no transverse motion as the lens is rotated. The sharpest image of the vertical lines (primary or **tangential focus**) and of the horizontal lines (secondary or **sagittal focus**) of the target is then located by the travelling microscope for vari-

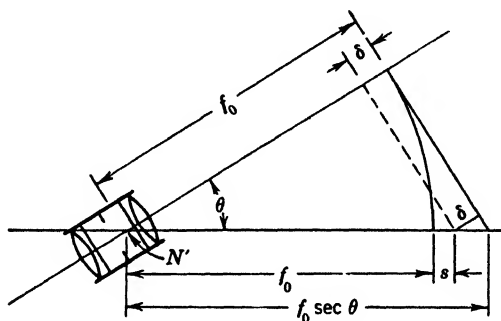


FIG. 6. Showing how the distance δ of an astigmatic image measured from the focal plane is related to the displacement s of the image measured along the optical bench.

ous angular rotations of the lens from 0° to $\pm 25^\circ$. The readings for $+5^\circ$ and -5° , and so on, should be nearly the same and should be averaged. These are the displacements s of the image as measured along a secondary axis of the lens. In order to compute the distances δ of the tangential T and the sagittal S images with respect to the focal plane of the lens, one uses the formula

$$\delta = [f_0 \sec \theta - (f_0 + s)] \cos \theta = f_0 - (f_0 + s) \cos \theta$$

which is readily derived from Fig. 6. The resulting values are plotted against θ to give the loci of the positions of the two astigmatic foci. The distance between the two curves indicates the amount of astigmatism for each value of θ .

If the two astigmatic images are close together, their mean position indicates the curvature of the image. This will be a section of some surface of revolution and should approximate the focal plane within the useful field of the lens.

Experiment 3. Photometry and Illuminometry

Part 1. Measurement of Luminous Intensity

The candle powers of three lamps are to be measured on a large photometer bench, using a Lummer-Brodhun cube for comparison of illuminations. Let E_s be the illuminance on one side of a magnesium oxide plate S , Fig. 7, due to the standard source of candle power I_s , and let E_x be the illuminance on the other side

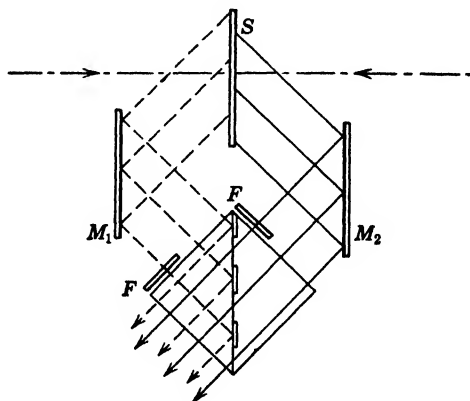


Fig. 7. Optical system of a Lummer-Brodhun photometer head. The light losses due to reflection at the plates F cause the trapezoids in the field of view to appear slightly darker than their surrounding areas. See Fig. 8.

due to the source to be tested, whose candle power is I_x . The illuminances are equal when

$$\frac{I_x}{d_x^2} = \frac{I_s}{d_s^2}$$

where d_x and d_s are the respective distances of the two sources from the screen S . The ratio of the candle powers

$$\frac{I_x}{I_s} = \left(\frac{d_x}{d_s} \right)^2 = R$$

may be measured directly on the scale of the photometer bench. Equality of illumination is judged by the matching of the two fields seen in the cube. In a Lummer cube, the hypotenuse faces make optical contact over part of their area, but are separated

elsewhere by etching away some of the glass. The light from one side of the magnesium oxide or carbonate screen is transmitted by the portions in contact, whereas that from the other side is totally reflected in the cube. Thus one sees a field like that shown in Fig. 8(a) or 8(b), where the numbers 1 and 2 indicate the respective portions of the two sides of S which are seen in the field of view. With the trapezoid field, it is customary to insert neutral filters F , Fig. 7, to make the trapezoids slightly darker than the surrounding field. If the sources have the same color temperature, the central division line disappears at balance, and the trapezoids show equal contrast with respect to the field which surrounds them. If the color temperature is different, the two halves of the field are complementary in color. One sees a pinkish trapezoid in a bluish field on one side and a bluish trapezoid in a pinkish field on the other. The photometer balance in this case is one at which there is equality of contrast on both sides, the trapezoid in each case being slightly darker than the field in which it appears. By matching contrasts instead of intensities it is possible to obtain reliable and reproducible settings even when the sources differ considerably in color.

Since the two sides of the screen S may not reflect equally, readings should be taken with the sides of the screen reversed. Moreover, the optical systems in the two paths may not transmit equally, usually because of different reflectivities at M_1 and M_2 . To eliminate this possibility, one should rotate the entire head through 180° and take readings in both positions. The ratio of the candle powers is then the geometric mean of the readings in the two positions of the head, as is shown below. Inserting the transmissivities T_1 and T_2 into the formula for the illumination balance, one finds for the first match

$$T_1 \frac{I_x}{d_x^2} = T_2 \frac{I_s}{d_s^2}$$

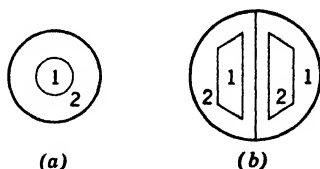


FIG. 8. Field of view in two types of Lummer-Brodhun photometer. The areas 1 are illuminated by light from one source, and the areas 2 are illuminated by the second. (a) The field of a Macbeth illuminometer. (b) The field in a large bench photometer.

so that

$$\frac{I_x}{I_s} = \frac{T_2}{T_1} \left(\frac{d_x}{d_s} \right)^2 = \frac{T_2}{T_1} R_1$$

Similarly, when the optical systems are reversed

$$\frac{I_x}{I_s} = \frac{T_1}{T_2} R_2$$

By taking the square root of the products of the respective sides of the last two equations, one finds that

$$I_x = I_s \sqrt{R_1 R_2}$$

The standard lamp is a properly aged incandescent lamp which has been calibrated at the National Bureau of Standards. The lamp should be placed at the same height as the screen *S* and with the etch marks on its bulb in line with the photometer axis, as directed in the calibration certificate. The voltage across its terminals is adjusted to the prescribed value. The lamp being measured is turned so that the identifying label is toward the screen and the "luminous center of gravity" is at the same elevation as the screen. The voltage across the terminals of each lamp being tested is adjusted to the values listed in the instruction sheet. One of the lamps is to be rotated through 90° and another set of readings taken in this position to illustrate the variation of candle power with direction.

Part 2. Measurement of Illumination and Brightness

Illuminance is defined as light flux per unit area and is commonly expressed in lumens per square foot, usually referred to as **foot-candles**. One may conveniently vary the illuminance on a surface by varying its distance relative to a 200-watt lamp. The high wattage is desirable so that the lamp need not be very close to the surface even for the highest scale-readings to be used, and consequently the illumination is more likely to be uniform. The illuminance is measured by means of a Weston foot-candle meter and by a Macbeth illuminometer. The Weston meter is of a direct-reading photovoltaic type. The only precautions necessary are that it be adjusted to read zero when the cell is covered, and that it be held so that the angle of incidence is the same, usually 0°, as at the illuminated surface.

In the Macbeth illuminometer a visual balance is made, with the aid of a Lummer cube, between the unknown illuminance and the illuminance provided by a standard lamp within the meter. The proper current through the standard lamp is determined with the aid of a reference standard which provides 3 ft-c when the current has its rated value. The standard is connected to one outlet on the control box, while the illuminometer is plugged in at the other side. Both are placed over the standard test plate which is provided. The current through either lamp is measured by the same milliammeter when the current switch is across the appropriate terminals. After determining the proper current to pass through the working standard lamp inside the illuminometer in order to obtain a match at the rated 3 ft-c of the reference standard when the illuminometer is set on this value, the standard is disconnected. This is important if the reference standard is to keep its rated calibration for a long time. Illuminances may then be read directly on the scale of the photometer, the current reading being kept at the value found by the use of the reference standard. The test plate is put where the illuminance is to be measured and is viewed at any angle within 30° of the normal. The distance from which one views the plate is immaterial so long as the test plate fills the rather small field of view of the instrument. The foot-candle scale runs from one to twenty-five, and the instrument is direct-reading in this range. In order to extend this range to either higher or to lower values, standard neutral filters are inserted into slots near the Lummer cube. The calibration of the filters should be checked by using an illuminance which may be read both with and without the filter or pair of filters. The ratio of the two readings gives the multiplying factor for the filter or pair of filters used.

The Weston meter is calibrated by comparing its readings with those of the Macbeth illuminometer at intervals of about 5 ft-c up to the highest reading on the Weston meter. The corrections to be added (algebraically) to the Weston meter readings should be graphed against its readings.

The Macbeth illuminometer is particularly useful in measuring brightness in reflected lumens per square foot, called **apparent foot-candles** or **foot-lamberts**. Since the standard reflecting plate has a diffuse reflectance of 80 per cent, an incident 3 ft-c illuminance, namely, 3 incident lumens per ft², will result in a dif-

fuse reflection of 2.4 lumens per ft². The current in the illuminometer lamp may be adjusted so that one obtains a match with the reference standard when the illuminometer scale reads 2.4 instead of 3. The instrument may then be used to measure directly the lumens per square foot diffusely reflected from any surface. To convert this reading to **millilamberts**, or 10^{-3} lumen per cm², one multiplies apparent foot-candles by a factor of

$$\frac{1000}{(30.48)^2} = 1.076$$

The brightness or luminous emittance of a surface will naturally depend on its illuminance as well as on its reflectance. It is instructive to make a few measurements using different materials. The diffuse reflectance is found by taking the ratio of the apparent foot-candles (reflected) to the actual foot-candles (incident).

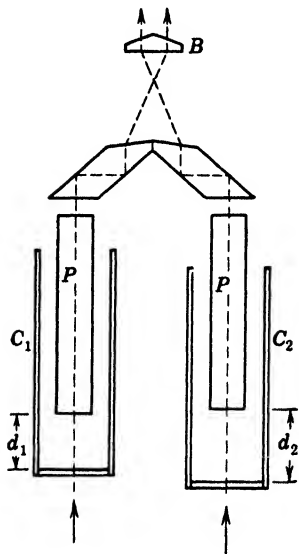


FIG. 9. Simplified sketch of the optical system of a Duboscq colorimeter. The cups holding the solutions may be racked up or down to vary the lengths of path d_1 and d_2 in the two solutions, P being stationary plungers of optical glass. The intensities of the transmitted light are compared with the aid of a biprism B which gives a sharply divided field.

Experiment 4. Colorimetry

In this experiment a Duboscq colorimeter, Fig. 9, is used to compare the concentrations of several series of solutions. According to **Lambert's law**, the intensity I transmitted through a thickness d of absorbing medium is given by the formula

$$I = I_0 e^{-ad}$$

where I_0 is the intensity before absorption, e is the base of natural logarithms, and a is the **absorption coefficient**. The value of a depends on the wave length of the light and on the concentration of the solution. In practically all cases, one finds that a is proportional to the concentration c so that $a = kc$, a relation known as **Beer's law**. Accordingly

$$I = I_0 e^{-kcd}$$

A colorimeter provides a ready means for varying the distance d of the light path through the two solutions to be compared (see instruction booklet for the colorimeter), and thus obtaining a measure of their relative concentration. The zero readings on the scales should first be checked by racking the cups, shown in Fig. 9, until the plungers P touch their bottom. If the reading is not zero, one may displace the vernier with the aid of a small screw driver. Secondly, the light intensities transmitted through the tubes must be matched, using the same value of d on both sides, about 15 or 20 mm, and with the same solution in both cups. To compare the concentrations of two solutions of the same substance, one of the concentrations usually being known, one matches the intensities of the light transmitted through both solutions, keeping the standard at a depth of 15 mm or 20 mm and varying the other. Then, since $I_1 = I_2$ at the match point, we have by the Lambert-Beer law

$$I_{01}e^{-k_1c_1d_1} = I_{02}e^{-k_2c_2d_2}$$

Since the intensities were originally balanced with k_1 , c_1 , d_1 respectively equal to k_2 , c_2 , d_2 , it follows that the intensities I_{01} and I_{02} were made equal and may be divided out of the above equation. By taking logarithms of both sides one finds that

$$k_1c_1d_1 = k_2c_2d_2$$

If Beer's law holds accurately for the substance being used, then $k_1 = k_2$ regardless of concentration, and consequently

$$\begin{aligned} \text{or} \quad c_1d_1 &= c_2d_2 \\ \frac{c_1}{c_2} &= \frac{d_2}{d_1} \end{aligned}$$

This is the working formula. The student should measure the concentrations of the solutions provided.

Experiment 5. Polarimetry

The objects of this experiment are (a) to study several half shades to determine the conditions under which angular settings are most accurate, (b) to measure the specific rotatory power of a

sample of sucrose, and (c) to use this constant in determining the amount of sugar in an "unknown" which the student himself prepares.

A simple polarimeter consists of a polarizer P , such as a Nicol prism, and an analyzer A mounted on a graduated circular scale to determine the direction of vibration of the polarized light, Fig. 10. If one sets the analyzer for extinction, with its principal section at right angles to that of the polarizer, and then inserts an optically active medium between the polarizer and analyzer, a change in angular setting will be necessary to restore extinction. Mono-

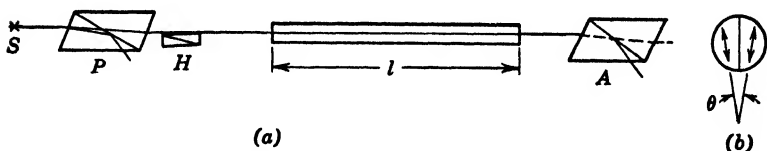


FIG. 10. (a) Passage of light through a polarimeter. P is the polarizer, H is a Lippich half-shade, and A is the analyzer. (b) The half-shade angle θ is the angle between the directions of vibration in the two halves of the field.

chromatic, usually sodium, light is used because the rotation depends on wave length. In order to increase the accuracy of the setting for extinction, it is customary to insert a half shade H just after the polarizer. There are several types of half shade, but they all have this in common: the field of view is divided into two (or three) areas in which the light vibrates in slightly different directions. The angle θ between the vibrations, Fig. 10(b), may be called the **half-shade angle**, and in many cases may be varied and set at some optimum value. To obtain the best results, the telescope should be focussed so that the boundary between the fields is as sharply defined as possible. The intensity in each field varies as the square of the cosine of the analyzer azimuth with a slight displacement of the intensity minima in the two or more fields. The intensities match four times in a complete revolution of the analyzer, twice at high intensity and twice at low intensity. The latter gives the most reproducible match, as one would expect from the **Weber-Fechner law**. This states that the smallest increment in intensity that can be detected is proportional to the intensity level. Hence one may detect smaller absolute changes in intensity at the lower intensity level.

Three half shades are provided in this experiment. These are to be compared as to accuracy of setting and, where the half-shade setting may be varied, the optimum value of θ is to be determined. The **Lippich half shade** consists of a narrow Nicol prism which covers part of the field (one half, or the central one third) as shown at *H* in Fig. 10. The vibrations passing through the Lippich prism are in its principal section, while those passing through the uncovered field are in the principal section of the polarizer. The half-shade angle may be changed by rotating the polarizer. The angle θ is first made zero by setting the polarizer so that the intensity over the entire field varies to zero concurrently as the analyzer is rotated. The repeatability of the readings is found by taking at least five settings of the analyzer and determining the average deviation from the mean, disregarding the sign of the deviation. Then the polarizer is rotated successively by 0.5, 1.0, 2.0, 4.0, and 8.0°, the procedure being repeated for each of these polarizer settings. The average deviation is plotted against the half-shade angle, and a smooth curve is drawn through the points. The small number of readings, five or so, may make the form of the curve uncertain. If this is found to be the case, additional readings should be taken; but five often proves to be sufficient to show the general trend and to determine the optimum value of θ .

The **Laurent plate** half shade consists of a half-wave plate of quartz or mica covering half the field. A half-wave plate gives a relative retardation of 180° between the principal vibration components, and the resultant of these is plane polarized, but in a different direction, as explained in Art. 12.11. The half-shade angle, measured relative to the zero setting which is found as before, is twice the angle through which the polarizer is rotated relative to the half shade. The student should find the optimum setting for the Laurent plate by taking a set of five analyzer readings at each of the following half-shade positions: 0, 0.5, 1.0, and 2.0°.

The **biquartz half shade** consists of two pieces of quartz cut perpendicularly to their optic axis, one of them being dextrorotatory and the other levorotatory. The thickness of quartz is such that light of the sodium wave length is rotated nearly 90° clockwise and counterclockwise respectively in the two halves of the field. Thus a suitable half-shade angle may be introduced, but this may not be varied by the experimenter since it depends on the thickness of the quartz. The accuracy of this half shade

may be compared with that of the other two by finding the average deviation from the mean of the same number of readings. It is interesting to try a source of white light with the biquartz. Because of the rotatory dispersion in the quartz, the two halves will in general have a different color. They may be matched by rotating the analyzer until the field is a uniform purple. The setting will correspond to the wave length which is rotated exactly 90° by the two quartz half plates. Thus measurements of rotation of the plane of polarization by an active material or solution using white light will not agree with the values obtained with sodium light.

The angle of rotation α by an active solution is proportional to the length of path l in the solution and to the concentration c , the constant of proportionality being called the specific rotatory power. Thus

$$\alpha = [\alpha]_{\lambda}^{\circ C} l c$$

where $[\alpha]$ is usually given in degrees per decimeter and grams of active substance per cubic centimeter of solution. The wave length is given as a subscript and the temperature as a superscript, since the specific rotatory power depends on both of these and it is necessary to specify them. It also varies slightly with the concentration, the amount being rather small, as seen by the formula for sucrose

$$[\alpha]_{5893}^{20^\circ C} = 66.462 + 0.870 c - 2.35 c^2$$

the concentration c being usually a small fraction. The specific rotatory power of sucrose also diminishes by 0.0184° per degree centigrade above $20^\circ C$. In this experiment the concentration effect will be by-passed by determining $[\alpha]$ for a standard value of c equal to 0.26 gm/cm^3 . This is common practice. Weigh an empty 25-ml bottle and then add about 6.5 gm of sucrose as determined by the increase in weight. For most accurate results, the sugar sample should be purified by recrystallization and dehydration. As this takes some time, and commercial sugar is quite pure, these operations will be dispensed with. Sufficient distilled water is added to the sucrose to make up 25 ml of solution. This is poured into one of the (clean) 2-decimeter tubes, and distilled water is poured into the other. With the tube of distilled water in the polarimeter all necessary adjustments of alignment and focus are made, using the half-shade setting found to give the best

results in the first part of the experiment. Then a set of at least five analyzer readings is taken to determine the "zero" reading. The tube filled with water is then replaced by the one containing the sugar solution, and the analyzer readings are repeated. The difference between the averages of the two sets of data gives α , from which $[\alpha]$ may be computed. The standard value of $[\alpha]$ may be obtained from the above formula, making a correction for temperature and concentration. The results of this experiment should be slightly below the accepted value principally because of the moisture in the sugar sample.

To illustrate the usefulness of polarimetry in analytical chemistry, make up an "unknown" containing about 0.40 gm/cm³ of sugar. Determine the concentration of this solution by measuring the rotation α and using the value of $[\alpha]$ found in the preceding part of the experiment. Since the same, somewhat impure, sample of sugar was used, the value found for c should check within about 1 per cent or the limit of accuracy of the balance used if this is not a particularly good one.

Experiment 6. Surface Testing by Interference

The standard **optical flats** used in this experiment are made of quartz or Pyrex glass to minimize warping due to temperature differences. Each has one surface which deviates from a plane by less than one quarter of a wave length. To obtain good optical contact between the flat and the surface to be tested, and to avoid scratching the flat, the surfaces should always be brushed by a clean camel's-hair brush before bringing them together. Further cleaning is unnecessary unless the surfaces are unusually dirty.

The surface to be tested is placed over the flat, or vice versa, according to relative size, the larger being always underneath. It is undesirable to have an appreciable amount of overhanging material because of the deformation it may produce. A viewer with a 45° reflecting plate is used, Fig. 11, so that the incident light may be approximately normal to the plate, and so that observations may be made from directly above. The source may be a sodium-vapor lamp or a mercury arc provided with filters to isolate the λ 5461 Å green line. The two surfaces are slid around until the number of fringes seen is a minimum. Sometimes it is

impossible to achieve this condition without recleansing the plates. A very slight pressure at some strategic spot may be helpful, but too much pressure should be avoided, as it warps the plates.

The fringes observed are fringes of constant thickness, discussed in Art. 10.8, the thickness of the air film being constant along any one band and changing by steps of half a wave length from one band to the next. Thus the view resembles a topographical contour map with the optical flat representing sea level in this case.

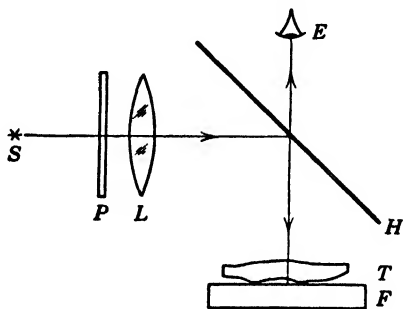


FIG. 11. Arrangement for testing the deviation from flatness of a surface *T* by interference. The fringes between *T* and the optical flat *F* are viewed through a 45° plate *H* which reflects monochromatic light down on the test object.

By studying the form and spacing of the bands, one may infer the form of the surface being tested. To determine which direction is "downhill," one may apply a slight pressure (e.g., with a pencil) at the location of any fringe. Its motion will be toward the greater thickness of air space or downhill on the surface to be tested, as viewed by an observer *facing* the surface. It may be noted that it is unnecessary to touch the plates to determine uphill or downhill; one may just look at the plates more obliquely. The fringe motion in this case is also toward greater thickness of air film, because at any one fringe

$$2t \cos i = p\lambda$$

remains constant. Consequently, as *i* increases and $\cos i$ decreases, the value of *t* at that fringe must increase to keep the product constant.

Several samples are provided, ranging from ordinary window glass to a sample of optical-flat quality. The surface of the window

glass is to be described only qualitatively, as it is extremely irregular. The telescope windows (beveled edges) and the gun-sight reticles are to be rated concave or convex to within some specified number of half waves or fringes in the directions of maximum and minimum curvature. The Newton's rings observed with these are sometimes approximately elliptical or hyperbolic, requiring two specifications of numbers of half waves to denote their principal curvatures.

A contour map is to be drawn for the sample of excellent-quality plate glass. This has roughly the shape of a col or saddle, Fig. 10.13. Draw "profile" sections through the directions of greatest positive and negative curvature.

Another flat placed on the standard flat yields a set of straight, equidistant, parallel fringes. As one reduces the number of fringes to a minimum, he may find a deviation of the fringes from straight lines, but the fringes are often so broad that the amount is not readily determined. After some patient, repeated trials, one may obtain an almost uniform intensity over the entire flat. This condition may mean that one "flat" is really somewhat convex, while the other is almost equally concave. By the use of a third flat, one may easily prove or disprove this assumption. All possible combinations of the three flats cannot fit unless they are all plane.

In testing a surface which is nearly flat, one finds it best to set it on the standard in such a way that about three or four fringes are observed, Fig. 12. The sagitta of the fringe connecting diametrically opposite points divided by the fringe separation gives the fraction of a half wave by which the flats deviate from each other across this diameter.

To measure the radius of curvature of a concave or convex surface one measures the diameters of several pairs of Newton's rings by means of a travelling microscope. An optical flat is used for the second surface. Since one can never be sure that the two surfaces are in contact at the center, the unknown separation e at this

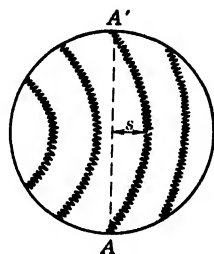


FIG. 12. If the surface being tested is nearly flat, one may determine any residual curvature by the sagitta s of an interference fringe which crosses it diametrically.

point, Fig. 13, is eliminated from the equations, using the diameters of two rings as explained in Art. 10.8. It is advisable to

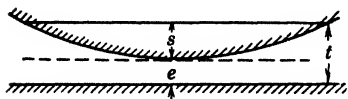


FIG. 13. Relation between the surface separation t , the sagitta s , and the central separation e between a convex spherical surface and a plane.

measure several pairs of rings, for example, the third and thirteenth, fourth and fourteenth, and fifth and fifteenth. The difference in order n is then equal to ten in each case. Rings of smaller order than three are usually too broad to measure with satisfactory results. The

value of the radius of curvature R is given by

$$R = \frac{d_2^2 - d_1^2}{4n\lambda}$$

where d_1 and d_2 are the diameters of the two rings which differ in order by n .

Experiment 7. Dispersion and the Prism Spectroscope

In a broad sense, the term **dispersion** refers to the variation of any quantity, such as index of refraction, with wave length or frequency. In spectroscopy, however, it often has a restricted meaning discussed in Art. 7.8, being defined by the expression

$$D \equiv \frac{d\delta}{d\lambda}$$

where δ is the angular deviation in the spectroscope measured in radians. To compute the dispersion and resolving power of a prism spectroscope, one must know how the index of refraction varies with wave length. This variation is conveniently expressed by either Cauchy's or Hartmann's formula, Art. 7.8, depending on the accuracy required.

In this experiment, the index of refraction of a prism will be measured for eight wave lengths from a mercury arc. The Cauchy constants n_0 and B are to be determined by finding the best straight line which can be passed through a graph of the points n plotted

against $1/\lambda^2$, preferably by applying the method of least squares to the numerical data. The applicability of Cauchy's formula is tested by using the values of n_0 and B to compute the values of n for the eight wave lengths used in this experiment, noting how the results deviate from the experimental values of n . The spectral lines to be used are indicated by letter symbols in Fig. 14, extra lines being drawn to facilitate identification. Since the yellow

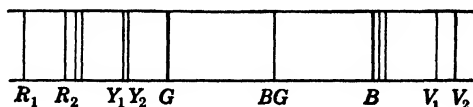


FIG. 14. The principal visible lines in the spectrum of a mercury arc. The letters indicate the colors of the lines. The table below gives their wave lengths.

lines are very close together, only one of them, Y_1 , is used in the measurement of n . The wave lengths of the eight lines to be used are, in angstrom units,

1. R_1	6908
2. R_2	6234
3. Y_1	5790
4. G	5461
5. BG	4916
6. B	4358
7. V_1	4078
8. V_2	4047

The red line R_1 is quite far from the other red lines and may not be seen by students with poor color vision in this part of the spectrum.

To determine the index of refraction one measures the angle of the prism and the angles of minimum deviation for each of the above lines. In order to obtain the required accuracy, it is first necessary to adjust the spectrometer properly and then make the settings and readings with care. It should be noted that the probable error of a set of measurements may be small if they are internally consistent, but the results will show a large systematic error if the instrument is not adjusted correctly.

Adjustment of Spectrometer

To facilitate the more exact adjustments which follow, one should first level the various horizontal parts of the spectrometer by inspection. A spirit level may be used but is not essential. The prism is removed, and the collimator and telescope, Fig. 7.16, are lined up in such a manner that their axes pass at least approximately through the vertical axis of the spectrometer. The refracting edge of the prism may be placed at the center of the prism table and used to indicate the spectrometer axis when sighting along the tubes.

A two-sided plane mirror is placed on the spectrometer table so that its plane is approximately parallel to the spectrometer axis. Such a mirror is best prepared by evaporating aluminum in a vacuum onto a flat plate of glass and cementing a protecting flat over the aluminum coating with Canada balsam. The eyepiece is focussed on the cross-hairs with the eyepiece as far out as possible, consistent with the obtaining of a sharp image. In this way the virtual image of the cross-hairs is placed at the far point of the observer's eye, this being at infinity if the eye is normal or if it is properly corrected by a spectacle lens. Under these conditions, accommodation is relaxed and consequently there is no eyestrain. A diffuse white light is placed opposite the opening in the side of the Gauss-type eyepiece. In this eyepiece some of the light is reflected through the telescope by a glass plate inclined at 45° and situated between the eye lens and the field lens of the eyepiece. When the mirror on the prism table is turned toward the telescope, this light is reflected back into the telescope. By focussing the telescope by means of the draw tube, one may see a reflected image of the cross-hairs. When this image shows no parallax with respect to the cross-hairs, the telescope is focussed for parallel rays.

The intersection of the reflected images of the cross-hairs is then brought into coincidence with the real cross-hairs by adjusting the levelling screws on the telescope and the prism table. Since both may be out of level, and by an unknown amount, it is best to make about one half of the adjustment with each of the two levelling screws as a first approximation. The double mirror is then rotated through 180° , and this adjustment is repeated until it is possible to bring the cross-hair images reflected from either side of the mirror into coincidence with the real cross-hairs by merely rotating the prism table. The double mirror is then par-

allel to the axis of rotation of the spectrometer, and the axis of the telescope is perpendicular to this axis. The latter object having been achieved, the mirror is removed and the collimator is focussed and levelled to conform with the axis of the telescope.

The prism is then placed on the prism table of the spectrometer in such an orientation that a rotation of any one of the levelling screws on the table will not change the level of a corresponding prism face. Another way of saying the same thing is to state that the prism faces should be respectively perpendicular to the lines joining pairs of the levelling screws. The reflected cross-hairs are then found by turning a polished face of the prism toward the telescope. This prism face is made parallel to the spectrometer axis by means of a prism-table levelling screw, bringing the reflected image into coincidence with the cross-hairs. It is important not to change the level or the focus of the telescope in this and the following adjustments. After both polished faces of the prism are made parallel to the axis of the spectrometer, the adjustment is completed.

The angle of the prism is most readily found from the readings of the scales and verniers of the spectrometer when first one prism face and then the other is set perpendicular to the telescope with the aid of the Gauss eyepiece. The difference between these readings differs from 180° by the angle A of the prism. Both verniers should be read, although it is necessary to read degrees on only one of the scales, since they are always 180° apart. Although it may be possible sometimes to obtain exactly the same vernier readings in some two positions on the circle, the difference between the readings on the two verniers will in general vary in a sinusoidal manner as one rotates the prism circle. The amplitude of this sine wave indicates the degree of departure of the center of the outer telescope circle from the axis about which the inner, prism circle rotates. Because of this eccentricity, any angle measured on one vernier will be a little high, while that measured on the other circle will be a little low. By averaging the readings on the two verniers one eliminates this source of error.

To measure the angles of minimum deviation, a mercury arc is placed behind the slit. The inner graduated circle is clamped, but the prism is left free to rotate, so that the readings are not affected by the rotation of the prism and depend only on the location of the telescope. Each spectrum line is seen to reverse its

direction of motion at the position of minimum deviation. It may be advantageous first to observe this effect without the telescope by looking into the prism from some little distance. Then the telescope is swung into place, and one examines the motion of the lines relative to the cross-hairs. The telescope setting is adjusted so that the intersection of the cross-hairs is exactly at the reversal point for each of the lines in turn. The two verniers are read at each of these positions of minimum deviation. The settings should be repeated several times and show no significant variation beyond the accuracy with which one can read the vernier. Finally, the prism is removed, and several settings and readings are made on the direct beam from the collimator. A more accurate procedure is to leave the prism in place, but to rotate it into its complementary position, where a second set of observations of minimum deviation are made on the other side of the direct beam. The differences between readings for each line are then equal to twice the angle of minimum deviation, giving roughly twice the precision of a direct measurement of δ . Another advantage is that the prism is still in adjustment after a complete set of readings so that any one of the first group of readings may be repeated and checked if some inconsistency is discovered.

The index of refraction of the prism is computed for each of the specified lines by means of the formula

$$n = \frac{\sin \frac{1}{2}(A + \delta)}{\sin \frac{1}{2}A}$$

The Cauchy constants are found from these results, and the dispersion and resolving power are computed for any specified wave length, such as that of the yellow lines at $\lambda_{av} = 5780 \text{ \AA}$. It may be difficult to obtain a good check with the dispersion found directly from the measured angular separation of the yellow lines because this separation is not very large compared with the error of setting and reading. If this is found to be the case, one of the yellow lines and the green line at $\lambda 5461 \text{ \AA}$ may be used. In this event, the average wave length of 5625 \AA should be used in the theoretical formula.

The theory of resolving power, Art. 7.8, shows that this quantity is proportional to the width of the beam leaving the prism. Thus, by using an auxiliary adjustable slit in front of the telescope, one

may reduce the aperture, and hence the resolving power, to any predetermined value such as that required to just separate the two yellow lines. This value is

$$r = \frac{\lambda}{\Delta\lambda} = \frac{5780}{21}$$

If the width of the auxiliary slit is w when this resolving power is obtained, the value for the whole prism R is found by the proportion

$$\frac{R}{r} = \frac{W}{w}$$

where W is the width of the beam leaving the unobstructed prism. The value of W is equal to the length of the second refracting face L multiplied by the cosine of the angle of emergence i' . The "experimental" value of R is then

$$R_{\text{ex}} = r \frac{L \cos i'}{w} = r \frac{L \cos \frac{1}{2}(A + \delta)}{w}$$

This should be compared with the "theoretical" value, Art. 7.8,

$$R_{\text{th}} = t \frac{dn}{d\lambda} \doteq \frac{2Bt}{\lambda^3}$$

The value of w should be measured several times with a travelling microscope. Settings of the auxiliary slit should be made at the exact point where the yellow band shows a barely perceptible shadow down its center, indicating the presence of a double maximum of intensity. The collimator slit should be as narrow as is consistent with obtaining enough light to make this observation. It may be necessary to clean the slit and to refocus the collimator before the line images are sharp and uniform. An agreement between the experimental and theoretical values of R of about 10 per cent is the best that one may expect in this part of the experiment. This is due to the indefiniteness in the setting of the auxiliary slit and also to the arbitrariness of the Rayleigh criterion for resolution on which the theory is based.

Experiment 8. Index of Refraction by Critical Angle

The index of refraction of a liquid, or of a solid which has at least one plane surface, may be conveniently measured by some form of refractometer. The first to be described here is a parallel plate refractometer devised by Pfund.¹ The lower surface of a plate of glass, Fig. 15, is coated with a white paint, onto which a small spot of light is focussed by means of a special light source. This consists of a small light bulb operated off a 6-volt transformer.

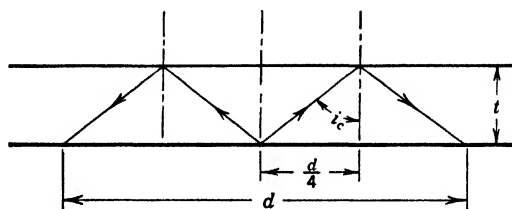


FIG. 15. Total reflection at the critical angle in a Pfund type of parallel plate refractometer. The diameter d is denoted by d_0 when there is air above the plate.

The light bulb is painted with an opaque black paint except for a small window about 2 mm square. The light from this window is focussed by a short-focus lens to form a small, intense spot of light. The light is diffusely reflected by the lower painted surface of the glass plate, so that it strikes the upper surface at a variety of angles. The rays striking at or beyond the critical angle are totally reflected back onto the painted surface, forming a ring of light with a sharp inner boundary. If t , Fig. 15, is the thickness of the glass plate and d_0 is the diameter of the inner boundary of the circle of light, the tangent of the critical angle of the glass plate is given by

$$\tan i_c = \frac{d_0}{4t}$$

Since

$$\sin i_c = \frac{1}{n_g}$$

¹ PFUND, *Journal of the Optical Society of America*, **21**, 182-186 (1931), or WOOD, *Physical Optics*, 3rd ed., Macmillan, 1934, pages 70-71.

where n_g is the index of the glass plate, one finds that

$$n_g = \frac{\sqrt{d_0^2 + 16t^2}}{d_0}$$

The diameter of the circle d may be measured by means of a scale ruled on the lower surface of the plate or with the aid of dividers and a separate scale. The thickness t is measured by means of micrometer calipers at several points on the plate where the paint has been removed.

If a layer of liquid is spread over the upper surface of the plate, one sees a second ring of larger diameter. The index of refraction of the liquid must be less than that of the glass plate, or there will be no total reflection at the glass-liquid interface. If d represents the diameter of this larger circle, then

$$\tan i_c = \frac{d}{4t}$$

and

$$\sin i_c = \frac{n}{n_g}$$

from which one finds that

$$n = n_g \sin i_c = n_g \frac{d}{\sqrt{d^2 + 16t^2}} = \frac{d}{d_0} \sqrt{\frac{d_0^2 + 16t^2}{d^2 + 16t^2}}$$

With some care, one may readily obtain measurements of n which are good to three decimals. This refractometer has been used in studying rates of slow chemical reactions, such as the drying of oils and varnishes.

A prism spectrometer is employed in the second method. The adjustments of the telescope and prism are carried out by the usual procedure, Experiment 7, except that the collimator is not used and need not be adjusted. Light from a broad source strikes one surface of the prism at various angles, as shown in Fig. 16. The bare mercury arc is a broad enough source, but the beginner may find it advantageous to place a sheet of ground glass between the arc and the prism. Of the three rays shown in the figure, ray c grazes the surface and enters the prism at the critical angle r_c . All other rays, such as a and b , are deviated clockwise to a greater extent, so that, if one points a telescope toward the prism in the

direction c' , he sees a field which has a bright left-hand portion sharply bounded from a dark right-hand portion. If heterochromatic light is used, there will be one such boundary for each wave length and a progressive overlapping of the bright areas to the left of each. With a mercury arc one can easily see three

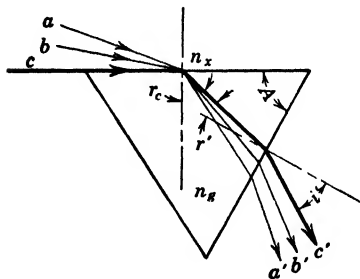


FIG. 16. Grazing rays such as c emerge at the smallest angle i' with the normal. On pointing a telescope in this direction, one sees the field sharply divided into a bright area on the left and a dark area on the right.

boundaries corresponding respectively to the principal wave lengths 5780 \AA , 5461 \AA , and 4358 \AA . One can determine the index of refraction of the prism and of any substance of lower index on the first face of the prism by measuring the angles i' and A . The angle A is measured by the method given in the preceding experiment. The angle i' is measured by setting the crosshairs of the telescope on the boundary formed by light of the wave length for which the index is desired, for example,

5461 \AA , and then setting the telescope normal to the prism face with the aid of the Gauss eyepiece reflection.

From the geometry of Fig. 16, one can see that

$$A = r' + r_c$$

By Snell's law

$$n_x = n_g \sin r_c$$

and

$$\sin i' = n_g \sin r'$$

from which one must eliminate r_c and r' to obtain the desired relation between n_x , A , and i' . Thus one finds that

$$n_x = \sin A \sqrt{n_g^2 - \sin^2 i'} - \cos A \sin i'$$

It should be noted that in some cases i' may be negative (on the opposite side of the normal), so that the second term must be added instead of subtracted. If the prism is surrounded by air, $n_x = 1$, one may solve for n_g , obtaining

$$n_g^2 = 1 + \left(\frac{\sin i' + \cos A}{\sin A} \right)^2$$

The index of the prism is determined by means of this equation and used in the preceding equation to find the indices of other substances.

The student should first measure the indices of refraction of two flint-glass prisms, using the wave length specified by the instructor. Then the indices of two liquids are measured. A few drops of the liquid are placed on a glass plate, and this is placed over the front surface of one of the flint prisms so that the liquid forms a film which wets most of the surface. The boundary giving the correct value of i' will be quite a few degrees toward the refracting edge of the prism and may be identified by the presence of interference bands within the bright part of the field. The last interference band is formed by light which just enters the prism at the critical angle. It is best not to have the film too thin, so that the interference bands are close together or perhaps imperceptible.

To measure the index of refraction of the glass plate, one proceeds as in the measurement of a liquid, except that he chooses a liquid whose index is greater than that of the glass plate. A consideration of the conditions for total reflection will show that only in this case may one have a grazing ray in the glass plate crossing over the film into the prism. A suitable liquid is monobromonaphthalin, $n = 1.658$, or perhaps methylene iodide, $n = 1.738$, if a still greater index is required. In all cases the index of refraction of the spectrometer prism selected must be greater than the index of the substance being measured. The boundaries are appreciably sharper if the surfaces are all thoroughly cleaned before use, and it is best if they are freshly polished surfaces, since old surfaces often acquire a tarnish which cannot be easily removed.

Experiment 9. Wave Lengths by the Plane Grating

The wave lengths of five lines in the mercury spectrum are to be measured by means of a transmission grating and a reflection grating. The spectrometer is adjusted according to the procedure given in Experiment 7. The transmission grating is then put on the table of the spectrometer, and its plane set parallel to the spectrometer axis with the aid of the Gauss eyepiece. The rulings

must also be parallel to the spectrometer axis. To make this adjustment, the slit is rotated into a horizontal position, or a piece of wire is fastened across the center of the slit with a bit of adhesive. Then, as the telescope is rotated, one observes if the slit or the wire images remain at the same height with respect to the intersection of the cross-hairs. If they do not, the grating is rotated in its own plane by turning the proper adjustment screw until all the spectra are at the same elevation as the central image. The slit is then returned to its usual vertical position.

In the discussion of the theory of the diffraction grating, Art. 11.6, it was shown that the principal maxima are located where

$$d(\sin i + \sin \theta) = m\lambda$$

the integer m indicating the order of the spectrum line, d being the interval between grooves. It is most convenient to set the transmission grating perpendicular to the collimator, so that $i = 0$, and

$$\lambda = \frac{d}{m} \sin \theta$$

The setting for zero angle of incidence is accomplished by first lining up the telescope and collimator, so that the zero-order slit image is on the cross-hairs. The Gauss eyepiece is then used to set the grating perpendicular to the telescope and therefore to the collimator. The grating is kept in this position while angle readings are made by setting the telescope on the spectrum lines.

The value of d is determined from observations made on the mercury green line, $\lambda\ 5460.74\ \text{\AA}$, in all the orders that one can find. This value of d is used in determining the wave lengths of the other four spectrum lines in any order, which may not be the same for all the lines. From the accuracy with which settings can be repeated, e.g., the average deviation, determine the probable accuracy of the wave length determinations and verify the measured wave lengths by comparison with tabulated values. Of course, if the adjustments are faulty, the error due to this cause may exceed the accidental error.

The transmission grating used may be a replica made by stripping a film of collodion or similar substance off a ruled grating. Such films are mounted between glass for protection. Their quality is variable because of some irregular shrinkage and im-

perfect mounting. The reflection grating is usually an original ruling on speculum metal, a mirror bronze of about 68 per cent copper and 32 per cent tin. The ruled surface is delicate and should never be touched. It is difficult to clean gratings without causing some damage. Instructions on the care of gratings may be found in Sawyer's *Experimental Spectroscopy*, pages 176-177.

The reflection grating is set on the prism table in place of the transmission grating, and the grating adjustments are repeated with it. The grating plane must be parallel to the spectrometer axis, and the rulings must be parallel to that axis, as before. The angle of incidence of the light on the reflection grating is arbitrarily set at some large value, such as about 70° . The orientation of the rulings is tested by comparing the height of the spectrum lines of various orders with the height of the direct reflection. As usual, a horizontal slit or a wire across the slit will be found helpful in making this adjustment.

The angle of incidence i is measured by observing the angle between the direct reflection of zero order and the normal to the grating as located by the Gauss eyepiece. The angles θ of the green line in all the orders are then measured with respect to this normal. The value of d is computed, using the grating formula. The wave lengths of five other lines are then measured, using the average value of d found with the green line. Any order may be selected. As before, some estimate of the probable accuracy should be made.

The resolving power of the grating may be found with the aid of an auxiliary slit. This slit is placed between the telescope and grating with its plane perpendicular to the telescope axis when this receives the yellow lines, Fig. 17. The slit is closed gradually until the two yellow lines just merge together into a band with a single maximum. The slit width is then measured by means of a travelling microscope. After repeating this several times, it is desirable to take also the same number of readings on widening the

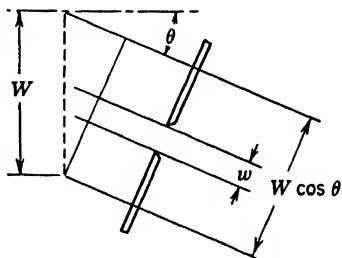


FIG. 17. Location of the auxiliary slit w used in reducing the resolving power of the grating to some known value r . The resolving power of the entire grating is proportionately greater than r in the ratio of $W \cos \theta$ to w .

slit until one can barely detect two maxima or a faint shadow down the center of the yellow band. These two sets of readings should be of the same order of magnitude and may be averaged. They give the aperture corresponding to a resolving power of $5780/21$, which will be denoted by the symbol r . Since the resolving power is proportional to the aperture, its value for the whole grating will be greater in the same ratio as

$$\frac{W \cos \theta}{w}$$

where W is the width of the ruled area, θ is the angle of diffraction, and w is the width of the auxiliary slit. The experimental value of R is then given by

$$R_{\text{ex}} = r \frac{W \cos \theta}{w}$$

and should be compared with the theoretical value

$$R_{\text{th}} = Nm$$

in which N is the total number of rulings. Since the two values cannot be expected to agree to better than 10 per cent, it is sufficiently accurate to measure W by means of a millimeter scale, from which N may be found by the relation $N = W/d$.

Experiment 10. Michelson Interferometer

The Michelson interferometer has been used in many kinds of measurement, some of them having been discussed along with the theory of the interferometer in Art. 10.9. Additional material may be found in Michelson's book, *Light Waves and Their Uses*, and in his *Studies in Optics*. In this experiment the interferometer is used to measure the ratio and difference between the two wave lengths of some close doublet, such as the sodium D lines, and to find the thickness of a sheet of transparent material such as mica.

Light from the source S , Fig. 10.16, is made approximately parallel by the lens L_1 and is divided into two beams of nearly equal amplitude by a thin aluminum coating on the back side of

the plane parallel plate M . The identical plate C is a "compensator" and serves to equalize the optical paths in the two arms for all wave lengths. The compensator may be rotated by means of a fine-pitched screw. The mirror M_1 moves on straight, parallel ways by turning a screw whose pitch is quite accurately 1 mm. The mirror M_2 is provided with adjustment screws to make its image in M either parallel to M_1 or at a small angle with M_1 . The metallized surfaces of the plates must never be touched, as they cannot be cleaned without more or less damage.

The interferometer is adjusted by first making the optical paths MM_1 and MM_2 as equal as possible with the aid of dividers or a scale. Looking into the interferometer, one sees several images of a wire placed between L_1 and M . The strongest of these images will be due to reflections originating at the aluminized surface of M . The screws controlling M_2 are adjusted until the strong images are in coincidence. It should then be possible to see interference bands. As M_1 is moved, these become curved. One can then refine the adjustment of M_2 so as to obtain a set of concentric circles, Haidinger's fringes, which do not expand or contract as the position of the eye is changed sideways and vertically.

As M_1 is moved, the rings expand or contract, with rings appearing or disappearing at the center, depending on the direction of motion. A motion of one half of a wave length completes one cycle of intensity variation at the center of the ring system. If two different wave lengths are emitted by the source, the intensity is simultaneously a minimum at the center when

$$2D_0 = p_1\lambda_1$$

and

$$2D_0 = p_2\lambda_2$$

where p_1 and p_2 are two integers, and $2D_0$ is the difference in path for coherent rays focussed at the center of the pattern. Because of the difference in wave length the minima will at first fall farther and farther from concurrence and then come closer together as D is increased. The visibility accordingly falls off and then increases to a second maximum. Figure 18 shows the successive positions of M_1 , giving dark-centered rings for two different wave lengths. They are seen to resemble a vernier and scale. Where the two positions agree, the rings have a high visibility; half way between,

the visibility is lowest. At the next agreement, or maximum visibility,

$$2(D_0 + d) = (p_1 + N)\lambda_1$$

$$2(D_0 + d) = (p_2 + N + 1)\lambda_2$$

where λ_2 is assumed to be the shorter wave length, as in the figure. The symbol d represents the displacement of M_1 between successive visibility maxima or minima, and N is the number of the longer-wave interference fringes passing by during the displacement d . By taking the difference between the second and the first of each pair of equations one finds that

$$2d = N\lambda_1$$

and

$$2d = (N + 1)\lambda_2$$

The ratio of the wave lengths is accordingly

$$\frac{\lambda_1}{\lambda_2} = \frac{N + 1}{N}$$

It is usually too tedious and not very accurate to determine N by direct count. If one knows one of the wave

lengths, the value of N may be computed from the measured value of d by

$$N = \frac{2d}{\lambda_1}$$

Since one can set on the visibility minima more accurately than on the maxima, the value of d is best determined by observing at least ten successive minima of visibility and obtaining the average interval by the method of differences. In this method one subtracts the first screw reading from the sixth, the second from the seventh, and so on, and divides the average of these differences by five.

An equation for the difference between the two wave lengths

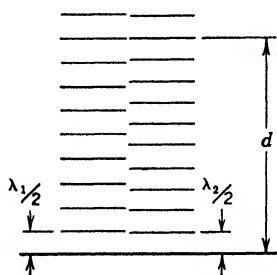


FIG. 18. Successive locations of the back mirror of a Michelson interferometer, giving central minima in the respective circular interference patterns due to two wave lengths λ_1 and λ_2 . The locations agree at intervals d which are positions of maximum visibility.

may be found by subtracting unity from both sides of the equation for the ratio of wave lengths, obtaining

$$\frac{\lambda_1 - \lambda_2}{\lambda_2} = \frac{1}{N}$$

Using the formula for N , one obtains

$$\lambda_1 - \lambda_2 = \frac{\lambda_1 \lambda_2}{2d} \doteq \frac{\lambda_{av}^2}{2d}$$

where λ_{av} is the average wave length, or the wave length as measured by some spectroscope which does not resolve the pair of lines. Thus one may obtain a value for a wave-length interval which is not easily resolved.

Michelson made noteworthy uses of this technique in the early days of spectroscopy in his pioneer studies of hyperfine structure and the Zeeman effect. He refined the technique by devising a method for analyzing visibility curves to obtain a determination of complex line structures and to measure the widths of spectral lines.

The thickness of a sheet of transparent material may be measured by determining the increase in optical path caused by introducing it into one arm of the interferometer. If μ_m is the index of refraction of the substance and μ_a is the index of refraction of the air which it displaces, the change in optical path is

$$2(\mu_m - \mu_a)t$$

where t is the thickness, and the factor two is introduced because the light passes through the material twice. If the optical retardation causes a displacement of m fringes of wave length λ , the thickness is given by

$$t = m \frac{\lambda}{2(\mu_m - \mu_a)}$$

To measure m one adjusts the interferometer for white light fringes. These are found when the path difference is a few half wave lengths. The central fringe of the pattern is achromatic, since it appears where the optical path difference is zero, and there

is accordingly the same phase difference of 180° for all wave lengths due to the external reflection at M . A beginner may have some difficulty in finding the white light fringes. A good procedure is first to adjust the interferometer for vertical fringes, using sodium light, and then to locate the place where the fringes are straightest and blackest. A rotation of the main screw by about one tenth of a millimeter in either direction should cause an equal drop in visibility and an equal and opposite curvature. The white light is placed behind the sodium flame, and its intensity is reduced until one can just see the sodium fringes faintly. This gives the observer something definite on which to focus his eyes, and he can readily see when the fringes become curved, indicating that he is outside the range where the white light fringes appear. By turning the main screw by small steps, one can then find the white light fringes without much difficulty. If the compensator is not exactly parallel to the 45° mirror M , the central fringe will not be achromatic, and one must proceed to rectify this fault.

Two wires may be used as sights onto which to set the achromatic fringe before inserting the thin plate to be measured. Fine adjustments are made with the compensator, and several readings are taken on its screw when the central fringe is on the wires. Then the thin plate is inserted in front of the mirror M_1 , taking care that the light passes through it perpendicularly. The white light fringes disappear over the portion of the field covered by the plate. They are brought back by increasing the slant of the compensator plate by means of its adjustment screw. Several readings are taken on this screw when the central fringe is again on the wires. The compensator scale is calibrated by taking readings on it for every fifth sodium fringe and plotting these readings against the fringe shift. This graph is used to determine the fringe displacement m from the compensator readings with and without the thin plate in place.

If the dispersive power of the plate is not the same as that of the compensator, the central fringe will not be perfectly achromatic, and one may have considerable difficulty in identifying the correct fringe. As a matter of fact, a constant error is introduced if the dispersive powers differ appreciably. The only way to eliminate this is to use a rotatable compensator plate of the same material as the one being measured. However, for a thin sheet of mica and a crown-glass compensator, the error is not very large and

will be ignored. If more accurate results are desired, the theory discussed in the first edition of Wood's *Physical Optics* should be consulted and applied.

Experiment 11. The Fabry-Perot Interferometer

This interferometer is more suited for the measurement of small differences in wave length than the Michelson interferometer. As explained in the discussion of the theory of this instrument, Art. 10.14, the rings are sharp so that one may resolve close pairs of lines instead of indirectly deducing their separation from the period of a visibility curve.

The instrument is adjusted by removing the collimating lens and observing the multiple images of the edges of the mercury-arc source. This may be diaphragmed down to make the multiple images more distinct. The images are first superposed by means of the capstan screws on the mounting of the forward mirror. When the collimating lens is replaced, one should be able to see some more or less distinct interference bands. They are made into complete concentric circles by further adjustment. The circles are then made to show no expansion or contraction as the eye is moved across the field horizontally and vertically. These final adjustments are made by means of the fine-adjustment screws which act through two helical springs on the front mirror support, causing a slight bending of the bar to which it is attached.

A special yellow filter is supported on the source side of the interferometer to isolate the yellow doublet 5770–5790. As one increases the separation of the plates, the rings are observed to coincide and separate into pairs periodically. One should note the direction of rotation which increases the separation of the mirrors and never rotate in the opposite direction without taking care to avoid jamming the mirrors together.

The coincidences of fringes correspond to the visibility maxima in the Michelson interferometer, whereas the bisecting alternations of the two ring systems correspond to the minima. The ratio of the wave lengths may be determined by measuring the period of this alternation of rings by setting on at least ten consecutive positions at which a bisecting alternation is observed. Then, as explained in the preceding experiment,

$$\frac{\lambda_1}{\lambda_2} = \frac{N+1}{N}$$

where

$$N = \frac{2d}{\lambda_1}$$

and

$$\lambda_1 - \lambda_2 = \frac{\lambda_{av}^2}{2d}$$

A more accurate method is to photograph the doubled ring pattern with a lens of accurately known focal length f and to measure the diameters d_1 and d_2 of rings of the same order in the two wave lengths. In this case

$$\frac{\Delta\lambda}{\lambda} = \frac{\Delta P}{P} = \frac{d_1^2 - d_2^2}{8f^2}$$

The quantity P is the order of interference at the center of the ring pattern and is not necessarily an integer. One usually knows the wave lengths with sufficient accuracy to be able to select a pair of rings having the same order of interference. If this is not the case, one may start with the plates very close together, practically in contact, and note as he separates them which rings belong to the same order. Other methods using several pairs of lines are described in Art. 10.14 and in the references given there. In these methods it is common practice to focus the complex ring system on a broadened slit of a spectrograph and simultaneously to photograph the spectroscopically separated diametrical sections of the ring patterns in the various wave lengths.

The atomic weight of deuterium may be determined from a measurement of the ratio of the wave lengths of the alpha lines of hydrogen and deuterium. It has been shown, Art. 17.2, that this ratio is

$$\frac{\lambda_{H\alpha}}{\lambda_{D\alpha}} = \frac{1 + \frac{m}{M_H}}{1 + \frac{m}{M_D}}$$

where m is the mass of the electron, and M_H and M_D are respectively the masses of the proton and the deuteron. The ratio

M_H/m is known to be 1836.6 from the specific charge e/m of the electron and the ratio of the Faraday constant of electrolysis to the atomic weight of hydrogen, the mass of the planetary electron being taken into consideration. If we let r represent the ratio M_H/m , then

$$\frac{\lambda_{H\alpha}}{\lambda_{D\alpha}} = \frac{1 + \frac{1}{r}}{1 + \frac{1}{r} \frac{M_H}{M_D}} = \frac{N + 1}{N}$$

The value of N is found from the period d of the alternations in the ring patterns by the formula

$$N = \frac{2d}{\lambda_{H\alpha}}$$

where the wave length of hydrogen alpha is taken to be 6562.8×10^{-8} cm. It will be found impossible to measure more than two or three of the successive alternations because the period is quite large in this case and the rings become small and broad as the mirrors are separated. Instead of readings on ten successive alternations of rings, the measurement of d should be repeated at least ten times over the same interval or over two intervals, if possible. A convenient formula for M_H/M_D is found by subtracting unity from both sides of the above equation for the ratio of wave lengths, giving

$$\frac{\frac{1}{r} \left(1 - \frac{M_H}{M_D} \right)}{1 + \frac{1}{r} \frac{M_H}{M_D}} = \frac{1}{N} = \frac{1 - \frac{M_H}{M_D}}{r + \frac{M_H}{M_D}} \doteq \frac{1}{r} \left(1 - \frac{M_H}{M_D} \right)$$

the latter approximation being permissible because r is much greater than M_H/M_D . Thus

$$\frac{M_H}{M_D} \doteq 1 - \frac{r}{N} = 1 - \frac{r\lambda_{H\alpha}}{2d}$$

The approximation is valid within the accuracy of this method for measuring ratios of wave lengths. If a method giving better accuracy than one part in 2000 is employed, the exact formula should

be solved for the ratio of atomic masses. If one takes the mass of the proton M_H to be 1.0076 amu, the mass of the deuteron is readily found. The addition of 0.00055 amu for the mass of the orbital electron is required to give the mass of the atom, but there is no point in making this refinement unless the experiment is performed quite accurately.

The light source used is of the type described by Harnwell¹ and illustrated in Fig. 19. The bulb B contains a small amount of heavy water mixed with an equal amount of distilled ordinary water. When the bulb B requires evacuation, the water must be

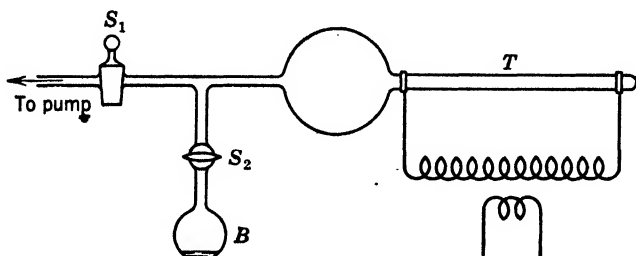


FIG. 19. Hydrogen and deuterium light source. Vapor is admitted into the tube T from the bulb B which contains a mixture of ordinary and heavy water. Pressure is adjusted by pumping or refilling.

frozen by a slush of dry ice and acetone or in finely pulverized dry ice. If the stopcock S_2 is properly greased and turned with caution, this does not have to be done very often. The source tube is excited by an oscillator of the Colpitts type, which is described in the article by Harnwell. The wire terminals are just wrapped around the outside at the ends of the tube T , which is viewed end on. The tube T is evacuated by means of a Cenco Hyvac pump. In order to keep air out of the tube and especially out of bulb B , the stopcocks are never opened until the pump reaches a low pressure equilibrium, as judged by its sound. Then S_1 may be fully opened, and, after equilibrium is restored, S_2 is slowly turned until a small amount of water vapor fills the tube, after which S_2 is immediately closed. The tuning condenser is adjusted for the setting which gives the most intense glow, and the pressure is adjusted until the source is as bright as possible. Then stopcock S_1 is closed to keep the pressure at the optimum value. It may be necessary

¹ HARNWELL, *American Physics Teacher*, **3**, 185 (1935).

to flush the tube with vapor several times before obtaining the best glow, which should be of a fiery red color.

A red filter is used to isolate the alpha lines. If the rings are photographed, a panchromatic film must be used. One may set

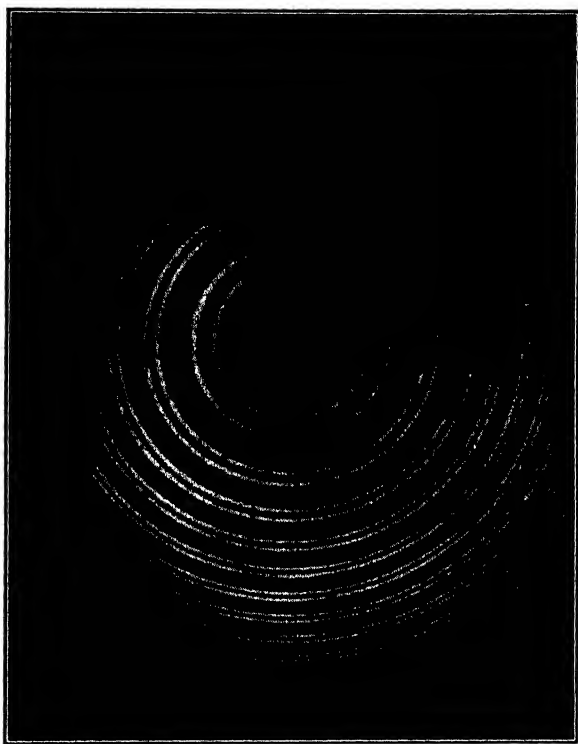


FIG. 20. Fabry-Perot interferometer rings using hydrogen and deuterium alpha radiation.

the aperture of the camera lens at $f/4$ and expose for about 10 min with the focussing scale set for "infinity." Figure 20 was obtained in this way.

Experiment 12. Spectrophotometry

Spectrophotometers consist of some form of spectroscope equipped with a device for the comparison of intensities in the spectrum of the colored light. A heterochromatic instrument

enables one to compare different wave lengths by means of a thermopile or photocell. A homochromatic instrument can be used to measure reflection or transmission coefficients at one wave length at a time. The instrument used in this experiment is the Hufner form of spectrophotometer manufactured by Hilger and is of the homochromatic type. The object of the experiment is to obtain spectrophotometric transmission curves for three samples of colored glass.

The Hufner rhomb, Fig. 21, situated in front of the slit, transmits two adjacent beams into the slit of a constant deviation spectro-

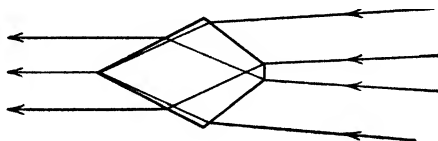


FIG. 21. Hufner rhomb. Note that the rhomb is proportioned so that the separated entering beams become closely adjacent emerging beams, and that the outer rays on one side become central rays on the other.

scope having a Pellin-Broca prism which can be rotated by means of a calibrated wave-length drum. One of the adjacent beams is passed through the sample, while the other is plane polarized by passing through a Nicol prism. Although the beams enter the slit adjacently, they are separated by a few millimeters at the sample holder (to the right in Fig. 21), so that the exact location of the lower edge of the sample is not critical. One should, however, take care that the frame of the sample holder does not obstruct any of the light entering the slit. By racking it up and down while looking into the instrument, one can easily see when it is too high or too low.

The wave-length scale is checked by placing a sodium light in front of the slit and setting the drum for a wave length of $589\text{ m}\mu$. The eyepiece is focussed on the shutters in its focal plane, and the left-hand shutter is set so that it bisects the image of the slit. If the prism must be turned to permit this adjustment, one must be careful to avoid heating it non-uniformly with the fingers or clamping it too tightly, as this would make the transmitted light elliptically polarized, so that perfect extinction could not be obtained. The left-hand shutter is to be kept fixed in position throughout the experiment, but the right-hand shutter is kept as far to the right as

possible, so long as one can obtain a uniform match of the upper and lower spectra along their common boundary when the sample is in place. If the absorption varies rapidly with wave length, the right-hand shutter must be moved to narrow the field, whereas in the flat regions of the absorption curve it may be kept a considerable distance away. It is possible to duplicate readings more easily if the upper and lower fields have a wide common boundary; hence this condition is preferred.

The light from both fields passes through an analyzing Nicol placed over the objective of the telescope. By rotating it, the intensity of the spectrum of the light passing through the sample is unchanged, but that of the comparison beam, which is polarized, may be reduced to match the former at any point in the spectrum. To make sure that the source is properly aligned, one removes the sample and sets the degree scale of the analyzer at zero. The upper and lower fields should match. If they do not, the source is moved or the condenser lens is moved until they do. A sliding neutral wedge is used to refine this setting.

According to the law of Malus, the intensity of the light transmitted through a polarizer and analyzer whose principal sections make an angle θ with each other is given by

$$I_1 = I_m \cos^2 \theta$$

where I_m is the maximum value of the intensity I . In the other half of the field, the intensity is proportional to the **transmissivity** T of the sample, so that

$$I_2 = TI_n$$

where I_n is the value of I_2 when $T = 1.00$, which is the case when the sample is removed. Since in the adjustment of the source one makes $I_1 = I_2$ when $\theta = 0$ and $T = 1$, it follows that

$$I_m = I_n$$

and therefore in general

$$T = \cos^2 \theta$$

when the two fields are matched. Readings of θ are to be made at intervals of 20 m μ from about 420 m μ to 740 m μ for the three samples provided. One should take several readings on both sides of the 90° (extinction) point, since the analyzer may read a little

high on one side and correspondingly low on the other. The values of the square of the cosine of the average angle θ are plotted against λ , giving a graph of T as a function of λ .

Some other quantities often used to describe absorbing substances are the following:

$$\text{Absorptivity: } A \equiv 1 - T = \sin^2 \theta$$

$$\text{Opacity: } O \equiv \frac{1}{T} = \sec^2 \theta$$

$$\text{Density: } D \equiv \log_{10} O = 2 \log_{10} \sec \theta$$

The values of A , O , and D are to be computed from the graphs for T at a wave length of $490 \text{ m}\mu$ for each of the three samples.

A supplementary or optional exercise is to determine the tristimulus values and the chromaticity for one of the samples by the method outlined in Chapter 20.

Experiment 13. Infrared Spectrometer

The optical system of the infrared spectrometer used in this experiment is shown in Fig. 22. Light from the source S is focussed on the entrance slit S_1 by a concave mirror M_1 . Partitions P_1 and P_2 are used to shield the thermopile T from the direct radiation from the source. A shutter is mounted on P_2 . The concave mirror M_2 reflects a parallel beam of light through the rock-salt prism P , where it is dispersed and then reflected by a plane **Wadsworth mirror** M_3 to a concave mirror M_4 . The mirror M_4 focusses the spectrum onto the plane of the exit slit S_2 . The mean wave length passing through the exit slit is indicated on a drum, the rotation of which turns the table on which P and M_3 are mounted, so that one may vary the wave length passing through S_2 in a known and continuous manner. Behind S_2 is a linear thermopile of the Coblenz type, which is connected to the Coblenz galvanometer G . This galvanometer is designed to have a high sensitivity and a low resistance, equal approximately to that of the thermopile. Since coils of the necessary low resistance and large number of turns would be too heavy for a delicate suspension, the coils are fixed, being imbedded in the soft-iron body of the galvanometer. Two groups of tiny magnets of opposite polarity are suspended on a fine quartz fiber in the magnetic fields of the two pairs of coils.

Connections are made so that the torques due to the fields of both pairs of coils are in the same sense, while those due to the earth's field and other stray fields are approximately compensated. However, in order to make the galvanometer quite independent of stray, variable fields, it must be magnetically shielded by about five concentric cylinders of soft iron. Even so, one may note some residual disturbances due to starting and stopping of elevators, nearby streetcars, and electric arcs in adjacent laboratories. The

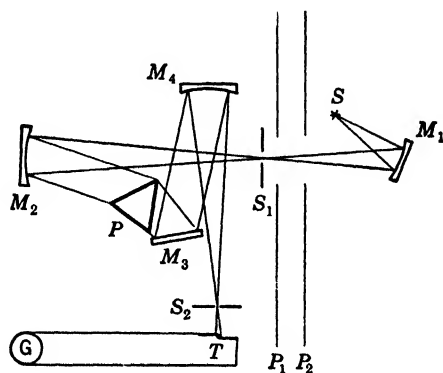


FIG. 22. Infrared spectrometer with a rock-salt prism P and a Wadsworth mirror M_3 for constant deviation. The wave length striking the thermopile T behind the exit slit S_2 is indicated on a graduated micrometer screw which rotates the table supporting P and M_3 .

carrier of the suspended magnets has a small mirror cemented to it, from which a fiducial line is reflected from a galvanometer lamp onto a translucent scale.

The wave-length drum should be checked by means of sodium light. The yellow lines at 0.589μ may be seen visually in an eyepiece placed behind the exit slit when the thermopile is moved aside. Care should be taken not to jar the thermopile in moving it, as this might crack some of the very thin wires or seals.

The purity of the spectrum depends on the width of the entrance slit S_1 , which should accordingly be quite narrow, consistent with sufficient intensity to obtain a sizable deflection. The thermopile may be evacuated to increase its sensitivity as much as sevenfold. The pressure required is beyond that which may be obtained by an ordinary rotary oil pump, since the mean free path must be made larger than the clearances in the thermopile chamber

to obtain a noticeable decrease in the thermal conductivity of the air.

The exit slit S_2 transmits a band of wave length $\Delta\lambda$, the magnitude of which is inversely proportional to the dispersion of the prism. The deflection is proportional to the total radiant energy flux ΔE contained in this band. Hence the energy distribution function

$$E_\lambda = \lim_{\Delta\lambda \rightarrow 0} \frac{\Delta E}{\Delta\lambda} \doteq \frac{\Delta E}{\Delta\lambda}$$

is proportional to the deflection d multiplied by the dispersion D . The dispersion is computed by the formula, Art. 7.8,

$$D = \left(1 - \frac{n^2}{4}\right)^{-1/2} \frac{dn}{d\lambda}$$

using data for n versus wave length from a table published by Coblentz. The computation may be accomplished most conveniently by graphing n against λ and finding the slope of the curve, $dn/d\lambda$, at several wave lengths. Using these slopes, the values of D are computed and plotted against λ , a smooth curve being drawn through the points. The values of D are then read off this curve for the wave lengths at which measurements have been made. The plot of $d \times D$ gives the energy distribution curve E_λ in arbitrary units. From the wave length of the maximum of this curve, one may find the approximate temperature of the source, using Wien's displacement law, which assumes that the source radiates as a black body.

The first part of this experiment consists of a determination of the energy distribution curve for the incandescent lamp provided. Since the filament of this lamp is a closely wound spiral, much of the radiation will come from the hotter portions between turns which approximate black-body conditions. The presence of the glass bulb around the filament limits the wave lengths transmitted to those shorter than about 4.5μ . If longer wave lengths are desired, a **Nernst glower** or a **Globar** heating element may be used. Readings should be taken at about $0.2\text{-}\mu$ intervals from 0.8μ to 4.4μ . The shutter should be closed between readings, and a zero reading taken at each point.

To measure the transmissivity of a substance, a plate of it is mounted on a carriage in front of the slit. By moving the car-

riage between stops, the sample may be inserted in the light path or removed from it. The ratio of the deflection with the sample in place to that with the sample removed gives the transmissivity. Except where the transmissivity changes rapidly with wave length, the width of the entrance slit may be increased, if necessary, to obtain conveniently large deflections. Readings should be taken at $0.2\text{-}\mu$ intervals from $0.8\text{ }\mu$ to $4.4\text{ }\mu$, and the transmissivity plotted graphically over this range for the samples provided.

Experiment 14. Concave Grating

The concave grating was invented by Henry Rowland in 1882. By ruling lines on a concave spherical mirror, at equidistant intervals along a chord, he was able to obtain spectra without the use of lenses. This type of grating has proved to be of great importance in the spectroscopy of the ultraviolet down into the soft x-ray region.

In Fig. 23, C represents the center of curvature of the grating and d is the distance between neighboring grooves, greatly enlarged. Let S represent a point on the slit, and P a focus of the diffracted rays at a principal maximum. The angles i and θ are the angles of incidence and diffraction, respectively, at the point

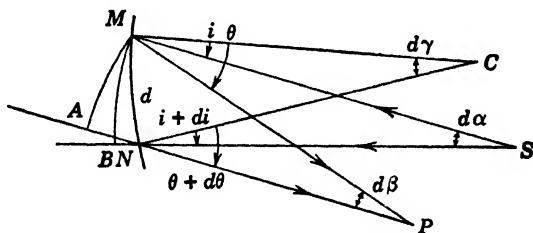


FIG. 23. Diffraction of light from S at a portion of a concave grating MN .

M on the grating. If MB is an arc whose center is at S and MA is an arc whose center is at P , the difference in optical path for a neighboring groove at N is

$$SMP - SNP = BN + AN = d(\sin i + \sin \theta)$$

There is a principal maximum at P if

$$d(\sin i + \sin \theta) = m\lambda$$

To find the locus of the points P for the various wave lengths or spectrum, one notes from the equality of vertical angles of two pairs of triangles in Fig. 23 that

$$i + d\gamma = i + di + d\alpha$$

or

$$d\gamma = di + d\alpha$$

and similarly

$$\theta + d\gamma = \theta + d\theta + d\beta$$

so that

$$d\gamma = d\theta + d\beta$$

Now,

$$d\gamma = \frac{d}{R}, \quad d\alpha = \frac{d \cos i}{s}, \quad d\beta = \frac{d \cos \theta}{p}$$

R being the radius of curvature of the grating, s the distance to the slit, and p the distance to the image from M . Thus

$$di = d\gamma - d\alpha = d \times \left(\frac{1}{R} - \frac{\cos i}{s} \right)$$

and

$$d\theta = d\gamma - d\beta = d \times \left(\frac{1}{R} - \frac{\cos \theta}{p} \right)$$

For a focus at P , the optical path $d(\sin i + \sin \theta)$ must remain constant over the grating, or, differentiating,

$$\cos i \, di + \cos \theta \, d\theta = 0$$

Hence, using the preceding expressions for di and $d\theta$, one finds that

$$\frac{\cos i}{R} - \frac{\cos^2 i}{s} + \frac{\cos \theta}{R} - \frac{\cos^2 \theta}{p} = 0$$

Now if S is on a circle of diameter $CM = R$, Fig. 24, then

$$s = R \cos i$$

and consequently

$$\frac{\cos \theta}{R} - \frac{\cos^2 \theta}{p} = 0$$

or

$$p = R \cos \theta$$

The phenomenon of the channelled spectrum has several applications. It may be used as a component of a color filter to greatly sharpen a transmission band. It may be used to measure the thickness of a very thin absorption cell or ultramicroscopic counting chamber,¹ or to find the thickness of the thin materials used as spacers to separate the two reflecting surfaces. It may also be used to calibrate a spectroscope independently of a large number of standard spectrum lines² or in a region where standard lines are not known. We shall measure t and calibrate a spectroscope.

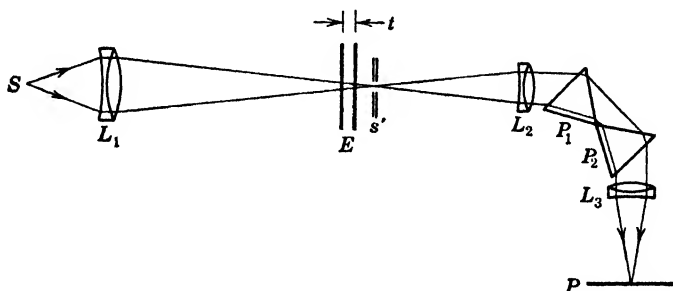


FIG. 26. Apparatus for observing a channelled spectrum due to selective interference in the Fabry-Perot etalon E . The source S is an incandescent lamp of the tungsten-ribbon type.

Light from a ribbon filament lamp at S , Fig. 26, is projected on the slit s' of a spectrograph by the lens L_1 . The spectrum formed by the prisms P_1 and P_2 with the aid of lenses L_2 and L_3 is projected on a ground-glass focussing screen at P . Light from a mercury arc is reflected into the lower half of the slit s' by means of a 45° prism (not shown). One thus obtains a continuous spectrum adjacent to the line spectrum of the mercury arc. The instrument is focussed on the mercury lines, and the slit is made just narrow enough to resolve easily the yellow doublet at 5770 – 5790 Å.

To obtain vertical, sharp interference bands, the etalon must be adjusted somewhat as a Fabry-Perot interferometer. The plates may not be very good optical flats, and in that case one merely centers the ring system as well as possible. The center of the ring system is placed opposite the uncovered part of the slit. By ob-

¹ PFUND, *Journal of the Optical Society of America*, **23**, 416 (1936).

² EDSER and BUTLER, *Philosophical Magazine*, **46** (1898); *Proceedings of the Physical Society (London)*, **15**, 207 (1898).

serving the fringes at P , while one moves or rotates E slightly, one may find a setting which gives narrow, vertical fringes. These are photographed, pages 421–429, on a panchromatic plate which is inserted with its emulsion side in the same plane as the ground surface of the ground glass. The panchromatic plate, being sensitive to all colors, including red, must be loaded and developed in complete darkness. It is advisable to practice inserting and removing a test plate from the plate holder before inserting the plate which is to be exposed. The time of exposure is of the order of 3

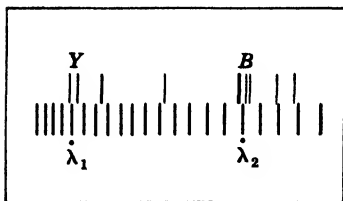


FIG. 27. Channelled spectrum with mercury-arc lines above. The etalon plate separation may be determined from the number of fringes between any two wave lengths.

sec to the light from the mercury arc, and about 10 min to the channelled spectrum. The plate is developed for about 5 min at 70°F or 3½ min at 80°F in a D-72 type of developer. After a rinse in water, it is placed in the fixing bath. The lights may be turned on after the plate has been in the fixing bath for at least 30 sec, but the fixing must be continued for 5 min after the plate appears cleared of all un-

exposed silver bromide. This will take about 10 to 15 min. The plate is washed in running water for 20 min, wiped with a viscose sponge, and left to dry. During all processing operations, the emulsion side must be turned upward, so that it does not stick to the tray and so that the solutions may have free access to the gelatin coating. Especially during development, it is advisable to rock the tray to make sure of uniform chemical action. If the plate is not panchromatic, a red safe-light may be used during the development.

The finished plate resembles Fig. 27. The wave lengths of the interference fringes in the vicinity of the yellow Y and the blue B lines are determined by linear interpolation between the Y pair and the B group respectively. Let these values be λ_1 and λ_2 . The order of interference at the first fringe is denoted by m_1 , while that at the second is $m_1 + n$, where n can be found by counting fringes, since it is known that the order increases by unity from any fringe to the next as one proceeds toward shorter wave lengths. Since

$$m_1 + n = \frac{2\mu t}{\lambda_2}$$

and

$$m_1 = \frac{2\mu t}{\lambda_1}$$

one finds, by subtraction and solving for t , that

$$t = \frac{n\lambda_1\lambda_2}{2\mu(\lambda_1 - \lambda_2)}$$

where μ , the index of the film, is unity if it is an air film.

Having found t , one can solve for m_1 by substituting into the equation for this quantity. Then, knowing m_1 and the fact that m increases by unity from one fringe to another toward shorter wave lengths, one can find the wave length of each fringe by

$$\lambda_x = \frac{2\mu t}{m_1 + x}$$

where x is the number of the fringe, counting from λ_1 , and is positive toward the shorter wave lengths and negative in the opposite direction.

One may obtain a calibration curve for the spectrograph by plotting the values of λ_x against their distance from any convenient standard line, such as the mercury green line at 5461 Å. Although the wave-length intervals from fringe to fringe decrease regularly toward the violet (the increasing dispersion makes it appear that the opposite is true), the wave-number interval remains constant and equal to $1/2\mu t \text{ cm}^{-1}$. The graph of wave number against x should accordingly be a straight line.

Experiment 16. Emission and Absorption in the Ultraviolet

The object is to compare the emission spectra of several sources and to determine the limits of transmission of various substances in the ultraviolet. Light from the source S , Fig. 28, is focussed on the slit s by the quartz condensing lens L_1 . This lens may be omitted if the source is large enough or may be placed close enough to the slit so that it fills the aperture of the collimator, or if the

source has sufficient intensity. The collimating lens L_2 renders the light parallel for some wave length which is preferably near the short-wave end of the spectral range. The **Cornu prism** Pr is made of dextro- and levorotatory quartz to compensate for the circular double refraction of its two halves, Art. 13.1. The prism disperses the light, which is then projected on the photographic plate P by the lens L_3 . In order that one may obtain a spectrum extending to about 2000 \AA , the lenses and prism must be made of crystalline quartz. However, the gelatin on a photographic plate absorbs below 2300 \AA , so that lines of shorter wave length are not photographed very well unless the plate is coated with a fluores-

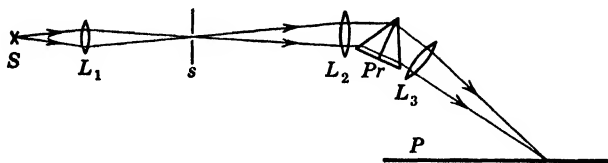


FIG. 28. Optical system of a quartz spectrograph.

cent oil. The oil must be removed before development by bathing the plate in ethylene chloride, acetone, or similar solvent. Since the lenses are rarely achromatized, the plate must be tilted as shown in Fig. 28 and also given a slight curvature by clamping it against a curved shoulder in the plate holder, so that all parts of the spectrum may be simultaneously in good focus.

The procedure for adjusting a prism spectrograph is discussed in detail by Sawyer.¹ Briefly, one must make sure that the prism is traversed by parallel light, the slit must be parallel to the refracting edge, the light rays should be parallel to a principal section, and the rays should pass through the prism symmetrically. Because of the variation of refractive index with wave length, it is impossible to satisfy these requirements completely. One can at best have them hold for one of the shorter wave lengths and the central part of the slit.

The spectrograph is usually in good enough adjustment so that one needs only to refine its focus. This is done by taking a series of spectrograms with different settings of the lens L_2 and selecting the one which gives the sharpest focus as determined by examining the plate with a magnifier.

¹ SAWYER, *Experimental Spectroscopy*, Prentice-Hall, 1944, pages 95-101.

The best location of the light source is found by opening the slit wide and placing the eye in the visible region of the spectrum, where one can see which portion of the optics is illuminated. The light source is moved until the light passes as centrally as possible through the lenses. The slit is then closed, care being taken not to damage the jaws by excessive force, and the zero reading is observed on the slit screw. The slit is then opened to at least the normal slit width and not more than seven times normal, Art. 7.8.

With each of the light sources aligned in turn as explained above, an exposure is taken of about 5-sec duration, racking the plate holder down about $7\frac{1}{2}$ mm between exposures. One of these sources is a mercury arc in a quartz envelope, the spectrum lines of which will be used for a wave-length calibration. Photographic procedures are discussed on pages 421-429, and some specific directions are given in Experiment 15. The mercury lines are readily identified on the finished plate by referring to a photograph of this spectrum, which is reproduced in Fig. 45 in the Appendix.

The limits of transmission of various kinds of glass are determined by placing the sample in front of the slit and using a carbon arc or an iron arc as a source of light. The limits are usually not sharply defined but nevertheless give approximate information as to the comparative range of ultraviolet transmission. Experiment 19 deals with quantitative measurement of transmission in the ultraviolet.

A wave-length calibration curve is obtained by graphing the wave lengths of identified mercury lines against the distance from some easily recognized line, such as λ 3341 Å. A line is drawn across the plate to indicate the location of this wave length on all spectra. This line can be drawn more accurately if a mercury spectrum is recorded at both the top and bottom of the plate. With the aid of dividers and scale, one measures the extent of each spectrum beyond this line. The corresponding wave length is determined with the aid of the calibration graph.

Experiment 17. Wave Lengths by Interpolation

The object of this experiment is to measure wave lengths on spectrograms by means of linear interpolation within an iron-arc spectrum and also by means of Hartmann's interpolation formula.

A large Littrow type of quartz spectrograph is used to obtain a sufficiently high dispersion for the many-lined iron-arc spectrum. The optical system is shown in Fig. 29. Light from the source *S* passes through the slit *s* and is totally reflected in the small quartz prism *P*₁. It is collimated by the lens *L*, dispersed by the prism *P*₂, and reflected back by a metallic backing on the prism *P*₂. The lens *L* then focusses the spectrum on a photographic plate *P*. The plate is in a carriage which may be racked up or down, so that about eleven spectra may be recorded at $7\frac{1}{2}$ -mm intervals. The

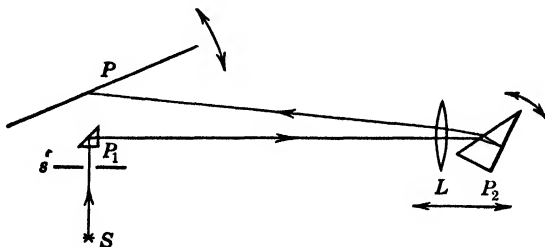


FIG. 29. Optical system of a Littrow type of quartz spectrograph. The short lines with arrows at each end indicate the three adjustments which must be made for any given spectral range.

exposure time is controlled by means of a shutter placed over the slit *s*. The spectral range covered on the plate is determined by three settings. They are (1) a setting of focus in which the lens and prism are moved along a short optical bench by means of a calibrated screw, (2) a rotation of the prism by means of a graduated micrometer screw, and (3) a setting of the tilt of the dark slide by means of a vernier and scale, the slide holder being hinged at one end.

Two sources are used, an iron arc of the Pfund type for the standard lines, and a Fuessner spark for the unknowns. The Pfund arc has iron electrodes, the upper of which is a quarter-inch rod projecting a short distance from a brass bushing which helps dissipate the heat. A bead of iron oxide is placed in the hollow of the lower iron electrode. The discharge to the bead of oxide raises it to incandescence and aids materially in stabilizing the arc. A direct current of about $3\frac{1}{2}$ amperes and 110 volts is used, 220 volts being preferable but not necessary. The source of the lines to be measured is a spark gap of the Fuessner type, Fig. 17.6, operated off a transformer giving 15,000 volts. A condenser is connected

across the transformer with an inductance and synchronous rotary spark gap in series with the spark source and connected across the condenser. The phase of the spark is adjustable by turning the Bakelite plate which supports the stationary electrodes. The spark is adjusted to give a uniform series of damped trains of oscillatory discharges across the source gap.

The source is aligned by opening the slit wide to minimize diffraction and rotating the prism to include a portion of the visible spectrum in the exit aperture. With the eye placed in this visible light and close to the location of *P*, one moves the spectrograph until the light passes centrally through the lens *L*. Of course, vertical adjustments are made by moving the source, which may be set to begin with at the same measured height above the table as the slit. Without disturbing the spark source, one then places the iron arc between the spark and the slit *s* and moves the arc until its light also passes centrally through *L*. Thus, after finishing the exposures with the iron arc, one may move it aside and have the spark source in the proper place for recording its spectra.

The zero reading of the slit is found by noting at which setting the light is just extinguished. The slit is then widened to about one or two times the normal slit width, which amounts to two or three small divisions on the drum in this case.

In order to check the adjustment and to improve it if necessary, it is advisable to make a series of exposures to the iron arc of about 3 sec each, the prism rotation being changed by one or two divisions between exposures. One should start below the recommended prism setting and continue an equal distance beyond. The plate is displaced vertically by $7\frac{1}{2}$ mm between exposures. Naturally, the dark slide must be closed before changing the prism setting. The plate is processed as outlined in Experiment 15, except that a red safe-light may be used if the plate is not panchromatic. The photographic process is discussed more completely on pages 421-429. The finished plate is examined with the aid of a magnifying glass, and the setting which gives the best focus is adopted for future use. The clear resolution of the iron triplet at 3100 \AA is a good criterion for satisfactory focussing.

The iron-arc spectrum is then recorded on a fresh plate. Without disturbing the plate in any way, one moves the iron arc out of line so that the light from the spark source may be photographed. The V-shaped diaphragm in front of the slit is displaced a few

millimeters so as to uncover a longer portion of the slit. An exposure is then made to the spark source, giving about 6 to 10 min of exposure time. One electrode of the spark source is an alloy of tin containing a small amount of tin and cadmium. The other electrode may be a piece of "spectroscopic" graphite rod.

The processed plate shows the tin alloy lines extending across the iron-arc spectrum. Being of greater length, they are easily located.

The wave lengths of the standard iron lines near the selected unknown lines from the tin spark are found with the aid of spectrum charts such as those published by Hilger or given in Brode's *Chemical Spectroscopy*. It is helpful to start by identifying a few lines at various points along the plate and plotting a crude graph of wave length against the distance from some easily located line. This graph may be used for an approximate determination of wave length, which indicates the region of the iron spectrum to be examined for standard lines. This being done, one identifies lines in the iron spectrum which are close to, and on each side of, each of the unknown lines. If d_1 and d_2 are the comparator readings for the two iron lines whose respective wave lengths are λ_1 and λ_2 , one may compute the wave length λ of the unknown line from its comparator reading d by the linear interpolation formula:

$$\lambda = \lambda_1 + (\lambda_2 - \lambda_1) \left(\frac{d - d_1}{d_2 - d_1} \right)$$

This should be carried out for the five or six lines selected by the instructor.

The wave length of one of these lines should be computed by the more accurate Hartmann formula. A third standard wave length λ_3 with its comparator reading d_3 is needed for this purpose. The three constants λ_0 , C , and d_0 in the (modified) Hartmann formula

$$\lambda = \lambda_0 + \frac{C}{d - d_0}$$

are determined by the simultaneous solution of the three equations obtained by substituting the standard λ, d pairs into the formula. The constant C is positive if the plate is measured so that d increases as the wave length decreases. If this is not the case, it is advisable to interchange d and d_0 in the denominator to keep C positive.

The standard lines should be within about 10 \AA of the unknown line to obtain its wave length within a few hundredths of an angstrom unit. Thus, for each unknown, one finds it necessary to recompute the constants. As pointed out by Hartmann, the value of λ_0 is practically constant for any wave-length range, but C and d_0 change considerably. One may use a single set of the three constants for a wide range of wave lengths, if their values are computed from standard wave lengths at the ends and near the center of the range, and if a correction curve is also used. This correction curve is found by computing the wave lengths of a large number of identified iron lines in various parts of the spectrum and plotting the difference between the known and computed wave length against the latter. A sinusoidal curve is obtained which crosses the axis at the wave lengths of the three standard lines used in computing the constants. For examples and further instructions, the student is referred to Sawyer's *Experimental Spectroscopy*, Chapter 9.

The computation of the constants in the Hartmann formula often leaves the student with some doubts as to the validity of the number of significant figures which must be retained. It is quite true that a slight change in the value of any one of the comparator readings d for any one of the known spectrum lines may lead to a hyperbola with considerably different parameters, so that, in this sense, the constants are not known very accurately, particularly if the three points are very close together. Nevertheless, the constants must be computed to a sufficient number of figures so that a substitution of the comparator readings of the three known lines gives computed wave lengths to the original number of significant figures. If this test fails, either there is a mistake in computation or an insufficient number of digits has been retained. It may seem paradoxical that, as the three points approach each other, the Hartmann constants become less certain, but the interpolated unknown wave lengths become more accurate.

It is accordingly instructive, and often convenient, to convert the Hartmann formula into a genuine interpolation formula by algebraically eliminating the constants C , λ_0 , and d_0 . One readily finds that

$$\lambda = \lambda_3 + \frac{1}{K + 1} (\lambda_1 - \lambda_3)$$

where

$$K = \left(\frac{d - d_1}{d_3 - d} \right) \left(\frac{d_3 - d_2}{d_2 - d_1} \right) \left(\frac{\lambda_1 - \lambda_2}{\lambda_2 - \lambda_3} \right)$$

The value of λ_3 is the smallest of the three known wave lengths. In this formula it is immaterial if the plate is put on the comparator so that the wave lengths decrease toward higher readings or vice versa. The second term on the right of the formula for λ is now a small correction term to be added to λ_3 , the term being smaller if the lines are closer together. Because of its smallness, it may be computed by means of a slide rule when the range of wave lengths is less than 10 Å and an accuracy of about ± 0.02 Å is sufficient. This treatment of Hartmann's formula arose from a discussion of the problem with Professor C. N. Wall.

Experiment 18. Quantitative Spectrochemical Analysis

The object is to determine the percentage of cadmium and bismuth in two samples of tin which is alloyed with these elements. The spectrograms for this experiment may be recorded on the same plate as in the preceding experiment. If this is not done, the same instructions must be followed down to the exposure of the photographic plate. After exposing the superimposed tin spark and iron-arc spectra, to help in identifying the alloy lines, the slit of the spectrograph is widened to about ten times normal and restored to its original length. This makes the lines wide enough so that the narrow exploring light of the microphotometer does not spill over the edges of the lines.

Seven tin alloys, with the concentration of cadmium and bismuth given in percentage by weight, are used to obtain the calibration curve. The concentration of bismuth and cadmium in the two unknowns marked X and XX is to be determined. A spectrogram is obtained for each pair of electrodes on a single plate, using an exposure of about 2 min and racking the plate down $7\frac{1}{2}$ mm between exposures. Both electrodes should be changed before each exposure, and the spark gap adjusted to a uniform length of 3 mm by means of a screw adjustment on the upper electrode.

After the exposure and processing are completed as in Experiment 15 (see also pages 421-429), the densities of the following pairs of lines are measured by means of a microphotometer

Sn 3143.86

Bi 3067.73

Sn 3655.54

Cd 3403.65

Sn 3655.54

Cd 3467.66

These lines may be readily located with the aid of the iron-spectrum chart or a rough dispersion curve drawn as explained in the preceding experiment. The Sn line in each pair serves as an **internal standard** to minimize errors due to fluctuations in the operation of the source. Other pairs may be used, but for best results the two lines of each pair must have the same excitation functions; when this is the case, they are called **homologous lines**.

The microphotometer in the author's laboratory is a modification of the Moll type made by Kipp and Zonen in Holland.¹ The optical system is indicated in Fig. 30. The lens L_1

concentrates the light on the slit s_1 , a reduced image of which is projected on the emulsion of the plate P by the reversed microscope objective L_2 . The transmitted light is projected by a similar lens L_3 through the slit s_2 onto a linear thermopile T , which is connected to the galvanometer G . The galvanometer is provided with a lamp and scale arrangement for observing its deflections. The lens L_4 , mirror M , and screen I are attached to an arm which enables one to move them as a unit into or out of the light beam of the densitometer. The lens L_4 is used to spread the light over a greater

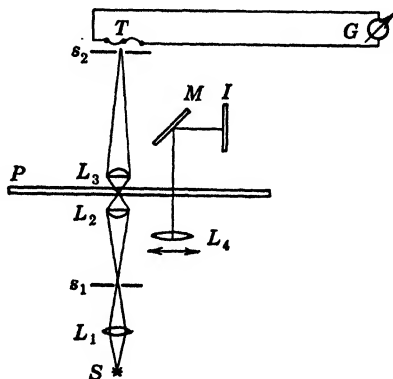


FIG. 30. Optical system of a modified Moll microdensitometer for measuring the photographic densities of spectral lines.

¹ The design of a much superior densitometer is given by Vincent and Sawyer in *Journal of the Optical Society of America*, **27**, 193 (1937). Other forms are described by Sawyer in *Experimental Spectroscopy*, pages 249-257.

area of P , of which the lens L_3 and M project an image on the screen I . At the same time, the mirror M serves to cut off the light beam to the thermopile, so that a zero reading d_0 of the galvanometer may be obtained. The image screen I has two marks inscribed on it, which indicate when the spectrum line to be photometered is in the proper position to intercept the light beam between L_2 and L_3 . The selected line should at first be placed a little to one side of these marks, where, on moving the combination L_4 , M , and I out of the light beam, as shown in Fig. 30, the clear plate deflection d_c is read on the galvanometer scale. The plate is then driven slowly by means of a screw so that the line image moves through the narrow light beam between L_2 and L_3 . The line reading d is taken when the galvanometer spot is closest to the d_0 reading. Then the light is cut off again by M , and the next line is moved into place. The density of each line is computed by the formula

$$D = \log_{10} \frac{1}{T} = \log_{10} \frac{d_0 - d_c}{d_0 - d}$$

T being the transmissivity of the photographed line image.

After all the line pairs are photometered in all spectra, one graphs the differences $D_{\text{Bi}} - D_{\text{Sn}}$ and $D_{\text{Cd}} - D_{\text{Sn}}$ against $\log_{10} C_{\text{Bi}}$ and $\log_{10} C_{\text{Cd}}$ respectively, C being the concentration. The graph should approximate a straight line if the plate has been properly exposed, i.e., if all densities are between about 0.3 and 1.3. With the aid of these graphs, established by the use of the known samples, one finds the concentrations of Bi and Cd in the unknowns, using their values of $D - D_{\text{Sn}}$.

In routine analyses of many samples one finds it advantageous to use a calibration graph of $\log_{10} I_x/I_{\text{Sn}}$ against $\log C$. The intensity ratio is determined with the aid of a plate calibration curve, giving $\log I$ as a function of D . The various methods for obtaining such intensity calibrations are discussed by Sawyer in his book *Experimental Spectroscopy*, pages 257–262. In general, standard samples are not used each time an unknown is to be analyzed, the $\log I_x/I$ versus $\log C$ curve being determined once and for all. After that, one measures the $\log I$ ratios with the aid of a plate calibration curve, which must, however, be determined for each spectrogram. The general procedure is outlined in Art. 17.4 and discussed more fully in Sawyer's *Experimental Spectroscopy*.

Experiment 19. Transmissivity by the Sector Photometer

The object is to measure the transmissivity of a sample of glass or other material in the ultraviolet. A Hilger sector photometer is used in combination with a small quartz spectrograph and a miniature hydrogen arc of the Urey type. The source is operated at about 3000 volts with a current of 200 to 300 ma. It emits a strong continuous spectrum¹ in the ultraviolet when viewed end

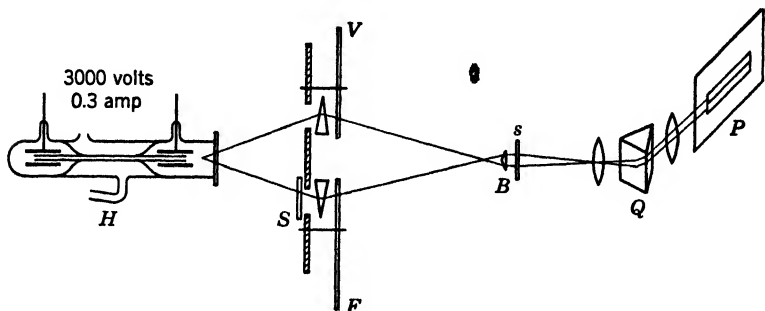


FIG. 31. Optical system of a Hilger sector photometer for measuring transmissivity in the ultraviolet region.

on. Water cooling is essential, the source being immersed in a copper tank through which tapwater is circulated. The light passes out through a quartz window, which is cemented to the Pyrex body of the tube. Light from the source *H*, Fig. 31, passes through two prismatic quartz lenses which focus the light onto a biprism *B* located in front of the slit of the spectrograph. With this arrangement two adjacent spectra are projected on the plate at *P*. The light in one of these has passed through the variable sector *V*, which is provided with a density scale. The other beam passes through the sample *S* and through a fixed sector *F*. The distances from the source to the sector and from the sector to the slit are made equal to the values stamped on a plate attached to the sector base.

It is essential to level and align the sector, the source, and the spectrograph according to the directions given in the pamphlet supplied by the manufacturer. With both sectors wide open, and

¹ HERZBERG, *Molecular Spectra and Molecular Structure*, Vol. I, pages 424–427.

the sample removed, the intensities of the two spectra are balanced as well as possible with the aid of a fluorescent screen. This may be a plate of uranium glass with one surface finely ground, or a plate of ground glass coated with moistened anthracene. If the lower spectrum is more intense along its entire length, the photometer should be lowered, and vice versa. If the lower spectrum is more intense only in the ultraviolet, the source must be lowered and the photometer lowered accordingly, and vice versa. When the two spectra are equalized as well as one can see, a test plate is exposed to check and to improve the adjustment. The general principles of the photographic process are discussed on pages 421-429, some specific directions for processing being given in Experiment 15. It is advisable to make nine exposures with the photometer at three different heights, and the source at three different heights for each of the photometer elevations. The heights giving spectra of equal intensity along their entire length are used throughout the experiment.

With the sample *S* in place one takes a series of spectrograms with the sectors rotating and the variable sector set at density values from 0.0 to 1.4 in steps of 0.1. It is advisable to increase the total time of exposure as the variable sector is closed, so that the match points will be at approximately equal densities. The following exposure table, based on the reciprocity law, may be followed:

<i>D</i>	<i>t</i> (seconds)	<i>D</i>	<i>t</i> (seconds)
0.0	2	0.8	12.6
0.1	2.5	0.9	15.9
0.2	3.2	1.0	20.0
0.3	4.0	1.1	25.2
0.4	5.0	1.2	31.7
0.5	6.3	1.3	39.9
0.6	8.0	1.4	50.2
0.7	10.0		

For the purpose of wave-length determination, a mercury-arc spectrum is registered at the top and bottom of the plate. After processing the plate by the usual procedure, a line is drawn across the plate through the 3341 Å mercury line. This serves as a reference line in all the spectra. A calibration curve is provided which gives wave lengths in terms of distances from this line, or a curve may be constructed as explained in Experiment 16.

With the aid of a low-power magnifier and a suitable scribe, one marks the points in each double spectrum at which the intensities are equal. At these wave lengths the sample has a density equal to the setting of the variable sector. These densities D may be graphed against wave length, or one may preferably graph the transmissivities T . These are computed from the densities with the aid of the formula:

$$T = \frac{1}{\log_{10}^{-1} D}$$

The extinction coefficient ϵ for an absorbing medium is defined by the formula

$$\frac{I}{I_0} = 10^{-\epsilon d (1 - R)^2}$$

where d is the thickness of absorber and R is the reflectivity at each face of the sample. From this formula one finds that

$$\log_{10} \frac{I_0}{I} = \epsilon d - 2 \log_{10} (1 - R)$$

so that

$$\epsilon = \frac{1}{d} \log_{10} \frac{I_0}{I} + \frac{2}{d} \log_{10} (1 - R)$$

In the absence of measurements of R , this quantity may be computed with the aid of Fresnel's formula for reflection.

One may question the validity of the variable sector as a means for obtaining an intensity (or density) scale, since one really varies the time of exposure of the slit to the full intensity of the source. In general, one may not assume that half as long a time of exposure, for example, produces the same change in photographic density as halving the intensity. In other words, the reciprocity law may not be relied upon. However, if one breaks up the exposure of both spectra into sufficiently many flashes, experiment shows that the resulting intermittency effect restores the validity of the reciprocity law as used in this experiment. This has been verified at the Bureau of Standards (*Technical Paper* 119) as well as by the manufacturers of the photometer. Sawyer warns, however, that the sector method does not give an accurate intensity

scale with all emulsions, in particular, with high-speed emulsions. Accordingly, they should be avoided when a sector is employed.

Experiment 20. Reflection of Polarized Light

As explained in Art. 12.3, the ratio between the amplitudes of the reflected R and incident E waves at an interface between transparent isotropic media is given by

$$\frac{R_s}{E_s} = - \frac{\sin (i - r)}{\sin (i + r)}$$

for the s components which vibrate perpendicularly to the plane of incidence, and

$$\frac{R_p}{E_p} = - \frac{\tan (i - r)}{\tan (i + r)}$$

for the p components in the plane of incidence. Hence

$$\frac{R_p}{R_s} = \frac{E_p \cos (i + r)}{E_s \cos (i - r)}$$

This relation is to be verified with the aid of polarizing attachments fitted to a laboratory spectrometer. The spectrometer is adjusted by the usual method, Experiment 7. The polarizer, which is equipped with a circular scale, is attached to the collimator at the lens end, and a polarizing eyepiece, also with a circular scale, is used instead of the usual Gauss eyepiece. To set the circular scales so that they are direct-reading, the polarizing Nicol is temporarily removed and replaced by a Wollaston double-image prism. The two images of the slit are in line with each other when the principal axis of the Wollaston prism is vertical. The eyepiece scale is set to read zero when the ray which vibrates vertically is extinguished. The polarizing Nicol is then replaced, and its scale set to read zero when the light is extinguished. The angular scale of the polarizer then reads directly the angle

$$\phi = \tan^{-1} \frac{E_p}{E_s}$$

and, at extinction, the analyzer reads

$$\theta = \tan^{-1} \frac{R_p}{R_s}$$

It is convenient to make $\phi = 45^\circ$ so that $E_p = E_s$.

A prism face is thoroughly cleaned with potassium hydroxide or with chromic acid and distilled water. This face is set so that its plane is in the spectrometer axis and is parallel to the collimator and telescope axes when they are in the same straight line so as to obtain grazing incidence. The ends of the slit are covered with opaque tape, leaving a small square central aperture which is illuminated with monochromatic light of 5461 \AA wave length obtained with a mercury arc and filter. It is convenient to set the spectrometer scale to read 180° when the light is incident grazingly on the prism face. The circular scale is clamped, but the prism and telescope are left free to rotate. The telescope reading in any position is then equal to twice the angle of incidence or $2i$. This angle is varied by steps of 10° which changes the angle i from 90° to the smallest angle obtainable before the telescope strikes the collimator, by steps of 5° . Corresponding to each angle i one observes the reflection from the prism face and reads the extinction angle θ on the analyzing Nicol scale. From the graph of θ against i one finds the Brewster angle as the point where θ passes through zero. The index of refraction of the prism is given by Brewster's law

$$\tan i_p = \mu$$

which should be verified by measuring μ by some other method.

Knowing μ , one can compute the angle of refraction r for each angle of incidence, and from these two angles the ratio

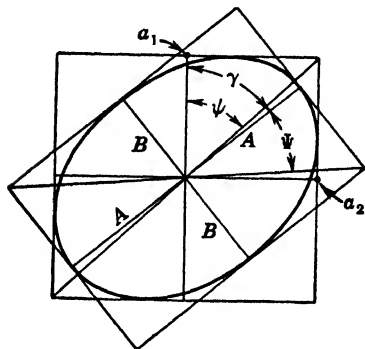
$$\frac{\cos(i + r)}{\cos(i - r)}$$

is calculated. The values of this ratio are compared with the corresponding measured values of $\tan \theta$, giving an experimental test of the Fresnel reflection formulas. The sign of the ratio of cosines may prove puzzling. It arises from the definition of the positive direction of the component R_p , which is preferable for nearly perpendicular incidence, but unnatural near grazing incidence. It may be found convenient in this experiment to redefine the positive direction of R_p and to insert a negative sign in front of the cosine ratio.

Experiment 21. Elliptically Polarized Light

By analyzing elliptically polarized light one may determine the thickness of a birefringent plate or measure the optical constants of a metal, depending on the manner in which the elliptically polarized light is produced.

When two coherent waves having the respective amplitudes a_1 and a_2 are vibrating at right angles to each other with a phase difference Δ , their resultant is elliptically polarized light. The semiaxes of the ellipse will be designated by A and B respectively, and the inclination of the A axis with respect to a_1 will be denoted by γ , Fig. 32. It is convenient to define the angles



$$\psi \equiv \tan^{-1} \frac{a_2}{a_1}$$

and

$$\Psi \equiv \tan^{-1} \frac{B}{A}$$

FIG. 32. Analysis of the constants of elliptically polarized light.

It is desired to determine a_2/a_1 and Δ by measurements of γ and Ψ with a quarter-wave plate compensator.

The spectrometer and its polarizing attachments are adjusted as in the preceding experiment with the device causing ellipticity removed during this process. A quarter-wave plate mounted on a graduated circle is attached in front of the telescope objective. Its circular scale is made to read zero when one of its principal vibration directions is in the same direction as the principal plane of the polarizer when this is at its zero setting. The light can then be extinguished by the (crossed) analyzer when its scale reads zero.

Elliptically polarized light for this experiment may be obtained by means of a thin plate of mica on the prism table or by reflection from a metal surface. This light is converted into plane polarized light when the principal vibration directions of the quarter-wave plate coincide with the axes A and B of the ellipse. There are two

positions of the compensator, differing by 90° , if it is exactly a quarter-wave plate, in which this may be achieved. The directions of the restored plane polarized vibrations in the two positions of the compensator differ by 2Ψ . The compensator scale gives the angles γ and $\gamma - 90$ in its two respective positions. Which of these angles is γ , the vertical slope of the major axis, may be determined by testing the ellipse with a plane analyzer, the compensator being removed. This angle is given approximately by the analyzer reading at minimum intensity, since it is then normal to the major axis. From the values of Ψ and γ one may compute the relative amplitudes and phase difference for the original components by means of the formulas

$$\cos 2\psi = \cos 2\Psi \cos 2\gamma$$

and

$$\tan \Delta = \frac{\tan 2\Psi}{\sin 2\gamma}$$

These formulas are proved in Schuster's *Optics*, pages 13-15.

If the elliptically polarized light was produced by a thin mica plate, the value of Δ is

$$\Delta = \frac{2\pi t}{\lambda} (\mu_2 - \mu_1)$$

Thus, if one knows the principal indices for the wave length used, one may compute the thickness t of the mica plate.

When the elliptically polarized light is obtained by reflection off a metal surface, it is best to set the polarizer at 45° and then proceed as above to analyze the resulting elliptical vibration. The angle of incidence i of the light should be measured by reading the telescope scale when a setting is made on the reflected light and when a setting is made on the direct light from the collimator. If these readings are T_1 and T_2 respectively, then

$$i = \frac{1}{2} [180^\circ - (T_2 - T_1)]$$

The index of refraction n of the metal is given approximately by

$$n = \sin i \tan i \frac{\cos 2\psi}{1 - \cos \Delta \sin 2\psi}$$

and the index of absorption is approximately

$$\kappa = -\sin \Delta \tan 2\psi$$

where Δ is the phase difference $\delta_p - \delta_s$.¹ The value of κ is related to the absorption coefficient α by the relation

$$\alpha = \frac{4\pi n\kappa}{\lambda_0}$$

where λ_0 is the wave length in air, and α is defined by

$$\frac{I}{I_0} = e^{-\alpha z}$$

The light is attenuated in the direction z , which is perpendicular to the surface. The propagation of light in metallic media is discussed in Chapter 14.

The value of $\Delta = \delta_p - \delta_s$ changes from 0 to $-\pi$ as the angle of incidence is increased from 0 to $\pi/2$. At some angle, called the **principal angle of incidence** \bar{i} , the value of Δ is $-\pi/2$. In this case the axes of the ellipse are in the directions of a_1 and a_2 , the angle γ being zero. One may accordingly set the compensator scale to read $\pi/2$ and vary the angle of incidence until the reflected light can be completely extinguished by just rotating the analyzer. The analyzer reading at extinction $\bar{\Psi}$ is the **principal azimuth**. From the values of the principal angle of incidence and of the principal azimuth, one readily obtains the optical constants n and κ of the metal by using the equations

$$\kappa = \tan 2\bar{\Psi}$$

and

$$n = \sin \bar{i} \tan \bar{i} \cos 2\bar{\Psi}$$

which are derived in Chapter 15.

The Babinet compensator is useful in measuring phase differences caused by double refraction or metallic reflection, Art. 12.12. The screw rotation s_0 required to displace the bands by one space, i.e., change the phase difference by 2π , gives the required

¹ Because of the direction of R_p , which is considered positive here (see Chapter 14), our Δ differs by π from the Δ used in Drude's *Theory of Optics*, page 363, for example, where similar formulas are derived.

information for translating a fringe shift of any magnitude s into the corresponding phase difference Δ . The relation is

$$\Delta = \frac{s}{s_0} 2\pi$$

The value of Δ determines merely the difference in phase between a_1 and a_2 without telling which is retarded with respect to the other. If a_2 is retarded, the elliptical vibration is clockwise and vice versa. To determine the sense of rotation of the ellipse, one notes the direction of the "fast" axis of the compensator and sets it on the major axis of the ellipse. If the principal plane of the analyzing Nicol when set for extinction is at an angle less than 90° and clockwise of the "fast" axis, the ellipse is rotating clockwise as seen by the observer, and vice versa. In this illustration the light would be called **left-hand** elliptically polarized light.

The Babinet compensator measures only the difference between the phases of the two components transmitted by an anisotropic plate. From this one obtains the relative retardation $(\mu_2 - \mu_1)t$, or the birefringence $(\mu_2 - \mu_1)$ when the retardation is divided by the thickness of the plate. The actual retardation $\mu_2 t$ and $\mu_1 t$ of either one of the rays is not obtained except by the use of some form of interferometer.

Experiment 22. Measurement of X-ray Wave Lengths

As Bragg has shown, the coherent wavelets scattered by the electrons in the three-dimensional periodic lattice of a crystal are in phase, Art. 11.11, when the following equation is satisfied

$$2d_n \sin \theta = n\lambda$$

The angle θ is the glancing angle, Fig. 33, of incidence and diffraction for a family of planes whose periodic spacing is d . The value of n gives the order of diffraction of the wave length λ by this family of planes.

If one takes into consideration the small refraction of x-rays by the crystal, the value of d is effectively different for the different orders, as explained in Art. 11.11. Thus, for the cleavage planes

of calcite, the effective values of d for successive values of n , as given by Siegbahn, are the following:

$$d_1 = 3029.040 \text{ XU}$$

$$d_2 = 3029.34$$

$$d_3 = 3029.40$$

$$d_4 = 3029.45$$

The unit in which these values are expressed is the **x-ray unit**, XU, introduced by Siegbahn, which is approximately 10^{-11} cm.

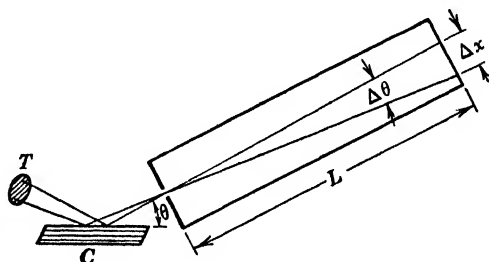


FIG. 33. Optical system of the Siegbahn tube spectrometer for photographing small regions in the spectra of x-rays of medium hardness.

Recent measurements of x-ray wave lengths by ruled gratings have shown that actually $1.00203 \text{ \AA} = 1000 \text{ XU}$, so that $1 \text{ XU} = 1.00203 \times 10^{-11} \text{ cm}$ (Birge, *American Journal of Science*, 1945). Although x-ray wave lengths are generally given on the Siegbahn scale, one must note the conversion factor to the cgs system in making accurate computations of densities or other properties of crystals and in applications of atomic theory.

For the measurement of wave lengths shorter than about 2000 XU, the Siegbahn tube spectrometer is most suitable. This employs a tubular camera of accurately measured length L with a slit at one end and the photographic plate at the other end, Fig. 33. The slit is placed near the crystal, which is mounted on a graduated circular scale. Radiation from the target T strikes the crystal at various angles because of the finite size of the focal spot and the crystal. For any point on the photographic plate P , however, the glancing angle θ has a definite value. One may simultaneously register a small portion of the spectrum, the range de-

pending on the location and size of the focal spot. To extend this range, one may rotate the camera and crystal in steps during the exposure, keeping their angular relation fixed. This is equivalent to moving T and is usually accomplished much more readily.

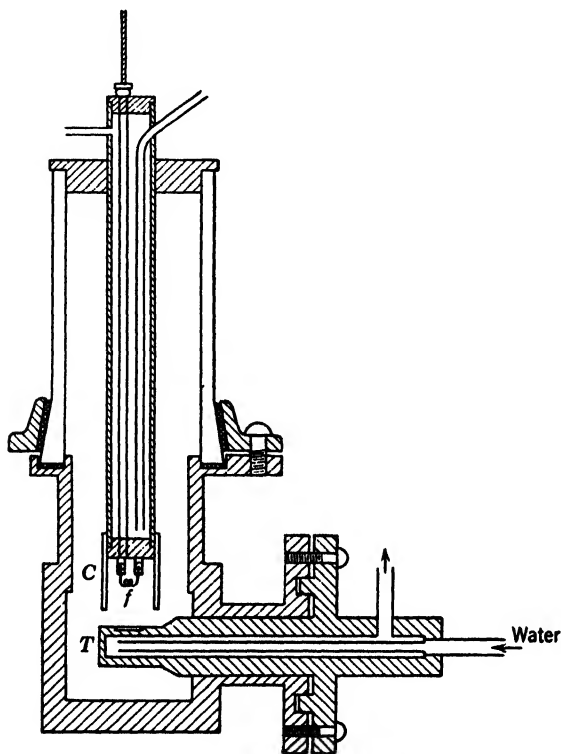


FIG. 34. A type of Siegbahn x-ray tube with rubber gasket seals for ease of disassembly. Electrons from the filament f are focussed on the target T by a cylindrical shield C . The x-rays from the focal spot pass from the tube through a window of aluminum foil.

The x-rays are generated by electron impact on a copper target, using a current of about 5–10 ma and a difference in potential of more than 30 kv. Since the target is water cooled, the tube is self-rectifying. The tube is evacuated by means of an oil diffusion pump with a suitable forepump. The pressure should be as low as possible to minimize the sputtering of tungsten from the filament onto the target. Such sputtering shortens the life of the

filament and introduces tungsten *L*-series lines, principally *L* alpha, which complicates the spectrum, Fig. 18.6. The lines from the body of the target itself are also weakened by absorption in the tungsten film.

A cross-section of the x-ray tube is shown in Fig. 34. The target body is a cube of copper into which one may dovetail plates

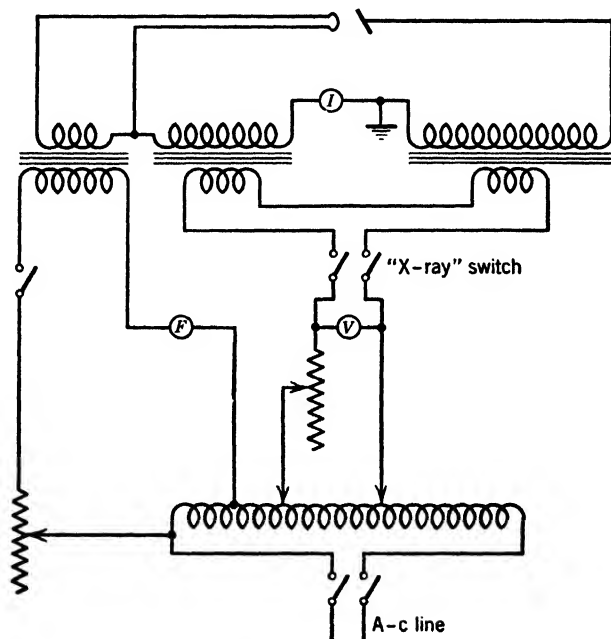


FIG. 35. Electrical circuit diagram showing transformers, switches, controls, and meters for operating a self-rectifying x-ray tube of the Coolidge, or hot-cathode, type.

of other metal. The spectra of compounds which are not readily dissociated may be obtained by rubbing the finely pulverized salt into scratches made with a sharp knife or razor blade on the face of the target. Figure 35 shows the wiring circuit for heating the filament and supplying the necessary high voltage to the tube. Before lighting the filament, the vacuum must be tested by a Pirani gauge or by momentarily applying a higher voltage across the tube than one intends to use. If no arc forms, the filament is turned on, without the high voltage, and pumping is continued to

remove the gases liberated by heat. After several minutes, the voltage is applied at a low value and raised gradually to the operating value as the tube outgasses. The complete process of outgassing should not require more than half an hour if the tube and vacuum system are clean and if there are no leaks.

The spectrometer is usually adjusted by the instructor or with his aid. If it is necessary to check the adjustments, a procedure similar to that outlined by Larsson, *Philosophical Magazine*, 3, 1136-1160 (1927), is followed. Even though the crystal and camera are adjusted, it is usually advisable to check the "zero" positions of both with the aid of a fluorescent screen. The crystal is set with its face parallel to the camera axis, and the slit is opened wide. The ends of the slit are blocked off, leaving an opening about 1 mm square. A fluorescent screen is placed in the plate position, and the camera and crystal are rotated together to find the position giving a maximum intensity, at which the angular scale of the camera circle is read. The crystal is rotated through 180° and the process repeated. The "zero" position for the camera is the average of the two scale readings obtained in this way. With the camera clamped at this reading, the crystal is rotated clockwise until the beam is noticeably reduced in intensity, and a reading is taken on the crystal circle. This is repeated for a counterclockwise rotation and then with the crystal rotated through about 180° . The average of all four readings on the crystal scale gives the "zero" setting for the crystal.

To measure the wave length of an x-ray line by Siegbahn's transposition method, which does not require the use of any standard lines, one makes two exposures on the same plate with the crystal in complementary reflecting positions as shown in Fig. 36. If the x-ray tube is fixed, so that the direction of the incident

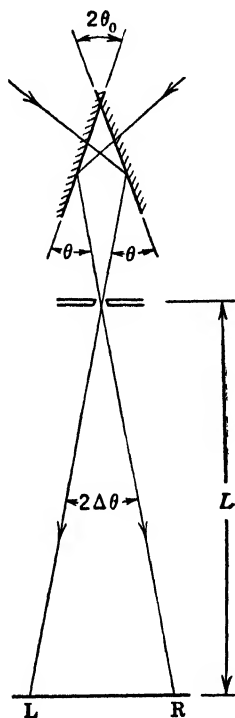


FIG. 36. Siegbahn's transposition method for "absolute" measurement of x-ray wave lengths.

rays is fixed, one obtains these by rotating the spectrometer and crystal together from $2\theta_0$ on the left of "zero" for the L exposure to $2\theta_0$ on the right of "zero" for the R exposure, the crystal being turned into its complementary position. The approximate value of $2\theta_0$ is computed by means of Bragg's formula and the approximate value of the wave length, this being generally known. θ_0 having been found, the crystal is rotated counterclockwise by θ_0 from its "zero," and the crystal and camera are rotated together through $2\theta_0$ clockwise for the first L setting. An exposure of about

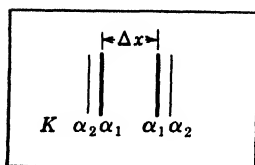


FIG. 37. Appearance of a spectrogram obtained by the transposition method, showing two pairs of the same K alpha lines recorded with complementary positions of the crystal.

1 hour at 5–10 ma and about 30 kv will be found suitable for the K alpha radiations from the iron group elements and for copper. After this exposure, the crystal is rotated clockwise through the angle $180^\circ + 2\theta_0$ to reflect x-rays in the opposite direction, the spectrometer tube being rotated through $4\theta_0$ in the counterclockwise direction for the R exposure. The crystal rotation is measured on the spectrometer circle to as high an accuracy as its calibration will allow. Siegbahn's precision spectrometer enables one to measure these angles to 0.1 second of

arc. The camera angle may be read by means of a rough scale, since it does not enter into the computations.

The photographic process is discussed on pages 421–429, some specific directions for developing and fixing being given in Experiment 15. The finished plate with its two pairs of K alpha lines should look like Fig. 37. The distance Δx is measured by means of a comparator, the value of $2\Delta\theta$ being obtained by means of the formula

$$2\Delta\theta = \frac{\Delta x}{L}$$

where L is the length of the camera. If this is not known, it may be measured by the method given by Larsson. The angle $2\theta_0$ is obtained from the crystal scale readings. Then, referring to Fig. 36, it is seen that

$$2\theta = 2\theta_0 + 2\Delta\theta$$

where θ is the actual Bragg angle. From its value one obtains the

wave length by means of Bragg's formula, using the value of d_n which corresponds to the order employed. If the L and R lines are found reversed on the spectrogram, the value of $2\Delta\theta$ must be subtracted from $2\theta_0$.

One may also measure any unknown line by reference to a nearby known line whose Bragg angle is computed. In this case, a single exposure is made, and the Bragg angle of the unknown line is found by adding or subtracting the angle between the two lines, which is determined by the formula

$$\Delta\theta = \frac{\Delta x}{L}$$

A suitable exercise consists of measuring the wave lengths of the tungsten L alpha lines, using the K alpha lines of copper as standards.

For wave lengths longer than 2 \AA , a vacuum spectrometer, Experiment 23, must be employed. Even then, one cannot measure wave lengths which are greater than $2d$, since for these $\sin \theta$ would have to be greater than unity. One can make some progress in the direction of longer wave lengths by using crystals with a large grating constant, but finally one must resort to ruled gratings at very oblique incidence.

Experiment 23. Wave Lengths of Soft X-rays

X-ray wave lengths are generally determined by measuring the Bragg angle θ in the formula

$$n\lambda = 2d_n \sin \theta$$

Siegbahn's transposition method for measuring θ without the knowledge of standard wave lengths is described in the preceding experiment. It is frequently more convenient, however, to determine θ by measuring the small angle $\Delta\theta$ between the unknown line and some standard neighboring line. Then

$$\theta = \theta_s \pm \Delta\theta$$

where θ_s is the Bragg angle for the standard line computed by Bragg's formula. If the standard line is not in the same order n as the unknown, one must be careful to use the correct value of d_n for each of the orders used, Experiment 22.

A plan view of the Siegbahn vacuum spectrometer is shown in Fig. 38, where T is the location of the target of the x-ray tube, s is the slit, C the crystal, and P the photographic plate. In this experiment one may rub the powdered, desiccated substance, whose spectrum is desired, into fine scratches in a metal (Al) plate which is dovetailed into the face of the target. As an exercise, one may measure the wave lengths of the chlorine $K\alpha$ lines in the radiation from sodium chloride. The third order of copper $K\alpha_1$ makes a

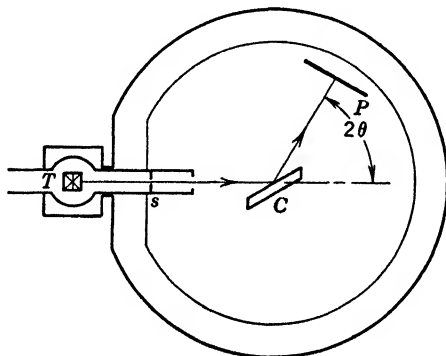


FIG. 38. Siegbahn's vacuum spectrometer. The tube T and the spectrograph are evacuated by a pumping system to extend the range of the spectrum to longer wave lengths.

suitable standard line whose θ_s is computed from the value $\lambda_s = 1537.40$ XU. Other standard wave lengths may be found listed in Siegbahn's *Spektroskopie der Röntgenstrahlen* or in Compton and Allison's *X-rays in Theory and Experiment*.

The radiation passing through the slit s falls on the crystal C and is diffracted to the photographic plate P . The distance sC must be equal to CP in order that one may rock the crystal during exposure without blurring the lines. The rocking is accomplished by making small arbitrary displacements of the crystal between $\pm 1^\circ$ of the computed angle θ_s for the standard line. A suitable tube current is 5–10 ma with a voltage of about 25–30 kv across the tube.

The spectrometer is usually adjusted by the instructor. If the student is required to adjust the instrument, the procedure in Lindh's *Röntgenspektroskopie* is to be followed.

The tube and spectrograph are evacuated by a pumping system consisting of a rotary two-stage oil pump backed by an oil dif-

fusion pump. The connection to the spectrograph by-passes the diffusion pump. A thin flake of graphite over the slit s separates the two evacuated chambers, permitting a lower pressure in the x-ray tube. The vacuum is checked by a Pirani gauge before operating the x-ray tube. The current and voltage must be increased gradually from rather low values to outgas the tube before the recommended current and voltage may be applied. It is necessary to use water-cooling for the target and body of the tube, and to cool the cathode assembly by circulating transformer oil through it.

An exposure of about 4 hours will be found ample to register all but the weakest lines. In order to distinguish between the unknown and the standard lines, it is convenient to interpose an opaque metal screen over part of the plate before exposing the Cu lines. The lengths of the lines will then be different in the two cases. The plate is processed by the method outlined in Experiment 15, except that one may use a red safe-light since the plate is not panchromatic.

The plate is measured on a Gaertner comparator, making at least ten settings on each line. The angle $\Delta\theta$ for each unknown line is computed by

$$\Delta\theta = \frac{\Delta x}{2R}$$

where Δx is the distance between the standard and unknown lines on the plate and R is the radius of the camera, i.e., $R = sC = CP$. In the author's spectrograph, $\log 2R = 2.37581$, distances Δx being measured in millimeters. The value of $\Delta\theta$ is to be added to θ , if the unknown line is on the longer wave-length side of the standard. One then uses Bragg's formula to compute the wave length of the unknown line or lines.

Experiment 24. Crystal Structure by the Powder Method

The powder method of crystal structure analysis was developed by Debye, Scherrer, and Hull. A fine powder of about 300 mesh is packed in a thin-walled soda-glass tube, or pressed into a rod, or stuck to a silk fiber. A multicrystalline wire of a metal may also

be employed. The sample is irradiated by monochromatic x-rays, which are usually obtained by passing the radiation from a molybdenum target through a zirconium filter or by passing copper radiation through a flake of nickel. The wave length, which is the K alpha characteristic radiation from the target, is 0.710 \AA in the first case, and 1.539 \AA in the second.

The beam is collimated by a slit system and strikes the crystalline powder located at the center of a cylindrical camera. The transmitted beam is allowed to pass out of the camera through a

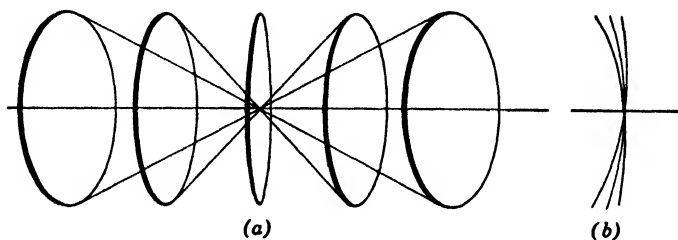


FIG. 39. (a) Cones of monochromatic x-rays diffracted by a powdered crystal. (b) Because of imperfect collimation, the traces on a photographic film are usually broadened. The broadening is least in the equatorial plane of the pattern.

hole cut in the film and through an opening in the far side of the camera. This opening is covered with black paper to exclude light. A convenient radius for the cylindrical sheet of film used to record the diffraction pattern is 57.28 mm, since each millimeter on the film will then correspond to an angle of 1° .

Each family of planes of spacing d_{hkl} gives rise to a diffraction halo of 4θ aperture, where θ is the Bragg angle for that family, Art. 11.11. The diffraction halos are somewhat as pictured in Fig. 39(a). The cylindrical film records only segments of these halos, these being sufficient to determine their angular apertures. Moreover, the slit character of the collimating system leads to an increased intensity, although there is a lack of definition at distances from the equatorial plane, since each ray through the slit is the origin of a vertically displaced arc, as shown in Fig. 39(b).

Before inserting the film, the slit system and specimen must be aligned in the beam from the x-ray tube with the aid of a small piece of fluorescent screen. A current of about 10 ma and a potential difference of about 30 kv will be found suitable. The film

is loaded into the camera with its central opening adjacent to the circular opening in the far side of the camera wall. After an exposure of about 6 hours, the film is developed by the usual method, Experiment 15. The diffraction pattern resembles that shown in

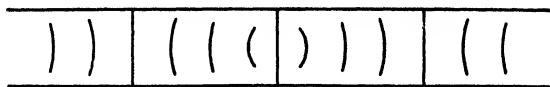


FIG. 40. General appearance of powder diffraction traces recorded on a cylindrical film.

Fig. 40. The appearance of the maxima depends on the absorption coefficient of the specimen. If this is quite high, there is a sharp outer edge because of the low penetration into the sample, as shown in Fig. 41. In this case, it is best to measure the distance D between the sharp outer edges of the corresponding maxima on the right and left, and to subtract the diameter d of the specimen in finding the angle 4θ by the formula

$$4\theta = \frac{D - d}{R}$$

On the other hand, if the specimen is very transparent to the radiation, one may measure the distance between the centers of the diffraction maxima, ignoring the quantity d .

From these measurements, one computes the values of $\sin \theta$, each of which is due to diffraction by some family of planes whose Miller indices are h , k , and l . If n is the order of diffraction, then for a cubic crystal these sines should be given by the formula, Art. 11.11,

$$\sin \theta = \frac{\lambda}{2a} \sqrt{(nh)^2 + (nk)^2 + (nl)^2}$$

where a is the lattice constant or length of the edge of each cubic unit of structure. From the formula given above, one finds that

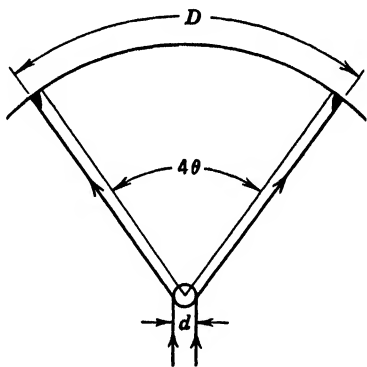


FIG. 41. The diameter d of a highly absorbing powdered sample increases the distance D between corresponding maxima.

$$\log \sin^2 \theta = \log \left(\frac{\lambda}{2a} \right)^2 + \log [(nh)^2 + (nk)^2 + (nl)^2]$$

The first term on the right is a constant for any pattern, while the second takes on values which are logarithms of integers obtained by giving nh , nk , and nl various integral values. This law enables one to find the proper indices for each diffraction maximum. The values of $\log \sin^2 \theta$ are laid off on a linear scale from the experimental data. Then a separate strip of paper is prepared with a plot of the logarithms of successive integers, using the same scale.

$\sin^2 \theta \propto$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Cubic (general)																		
Body - centered																		
Face - centered																		

FIG. 42. Comparison of the powder diffraction patterns for three types of cubic crystal lattice. The absence of certain maxima enables one to distinguish between them.

By sliding the second plot along the first until it fits the observations, with some gaps usually, one may infer the values of nh , nk , and nl from the values of the sum of their squares. Ambiguities will only occur with the higher numbers. One should note whether any maxima are missing or are abnormally weak because of destructive interference between atoms in special locations in the unit cell. For example, a face-centered cubic structure will not give certain maxima, and similarly for a body-centered structure. In the general cubic case, all maxima are present with a few exceptions such as 7 and 15, Fig. 42. A study of the relative intensities of the diffraction maxima enables one to determine the crystal group by methods which will not be discussed here.

In this experiment one generally analyzes the structure of a metal wire which is either face-centered or body-centered cubic. The diffraction pattern shows clearly which it is by comparison with Fig. 42. After the maxima are properly indexed with the aid of the logarithmic plot, the lattice constant a is computed, using each of the diffraction maxima, the results being averaged. If the radiation is not properly filtered, the diffraction pattern will consist of a superposition of two or more patterns which are displaced by a constant amount on the logarithmic plot because of the change

in wave length. These patterns may usually be disentangled without much trouble, since, if one knows the wave lengths, he may compute the amounts, $\Delta(2 \log \lambda/2a)$, by which the patterns are displaced. If one uses unfiltered radiation from a copper target, he may expect a weak pattern because of wave length 1.39 \AA , Cu $K\beta$, and perhaps because of 1.42 \AA , W $L\alpha$. The tungsten sputters onto the target in a Coolidge type of tube and finally obscures the copper lines as a result of absorption. The sputtering proceeds at a more rapid rate if the vacuum is poor. In order to avoid this difficulty, some experimenters prefer to use the old type of ion tube in spite of the problem of maintaining exactly the right degree of evacuation for smooth operation. Many ingenious devices have been developed to control automatically the pressure in such tubes; for example, see *Applied X-rays* by Clark.

Clark's book may also be referred to for information in regard to methods for analyzing crystals with a more complex structure than those discussed in this experiment. Another good reference is Buerger's *X-ray Crystallography*.

The Photographic Process

One advantage of photographic recording of spectra and diffraction patterns is that the process is integrating in nature, allowing one to record very weak radiations by giving a suitably long exposure. Another advantage is that a permanent record is obtained which may be studied at leisure and referred to later for details which may have not been noticed or thought important at first. Moreover, photographic plates are sensitive to a very wide range of wave lengths which extend far beyond the limits of vision. Their principal disadvantage in quantitative photometry is that photographic density depends on so many factors that it is very difficult to reproduce measurements exactly. It is often more accurate to measure the intensity of a spectral radiation directly by means of an electron multiplier tube than to derive the intensity indirectly from the recorded density of a photographic image. Hence for some purposes the direct measurement is preferable, although it is selective in the sense that one must choose the particular lines which he considers desirable to measure. For the exploration of an unfamiliar spectrum and for the measurement of wave lengths, the photographic method is particularly useful.

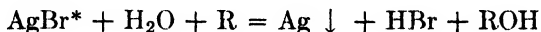
The photographic plate consists of a layer of gelatin in which a mixture of small silver halide crystals are imbedded. This so-called **emulsion** is supported on a transparent plate of glass or on a film of some cellulose ester. Gelatin has the remarkable property that it swells in aqueous solutions and permits the diffusion of chemicals to and from the silver halide grains while holding them in place. After drying, the gelatin returns to its original thickness and toughness. Of the silver halides used, the bromide has the greatest sensitivity and the greatest spectral range over which it is sensitive to light. Consequently it is commonly used in a practically pure state in negative emulsions, with only a small admixture of the iodide. The grain size covers a typical distribution curve with a size of 0.1 to 0.2μ predominating. The silver bromide is most sensitive to light of about 4800 \AA , the silver chloride to 3800 \AA , and the silver iodide to 4400 \AA . The sensitivities drop rapidly toward longer wave lengths but decrease only slowly toward shorter wave lengths.

It is well known that silver halides darken on exposure to light, but this action requires a relatively long exposure. A much greater sensitivity to light depends on the formation of an invisible **latent image** of silver salt, which is reducible to metallic silver by the action of a suitable **developer**. The formation of the latent image requires the presence of a small amount of sulphide as an impurity in the silver halide. The amount is now controlled in the manufacturing process. The sulphide generally comes from a small amount of mustard oil in the gelatin. The sulphur atoms provide localized energy levels in the halide crystal which serve as traps for electrons which have been raised to the photoconduction levels by the absorption of light. The negative charges accumulating at these centers then attract some wandering silver ions, forming a nucleus at which reduction of the grain may begin by the reducing action of the developing agent. For further details the original papers¹ should be consulted, the third paper listed below being a good introductory survey article.

The development of the latent image consists of its reduction to a granular deposit of metallic silver by means of a suitable reducing agent. Reduction is followed by the auxiliary processes of fixa-

¹ GURNEY and MOTT, *Proceedings of the Royal Society*, **164A**, 151 (1938); WEBB, *Journal of the Optical Society of America*, **26**, 367 (1936); WEBB, *Journal of Applied Physics*, **11**, 18 (1940).

tion, hardening, washing, and drying. Reduction is defined as the release of a metal from a combination with a non-metal, or, more generally, as a decrease in effective valence. If the symbol R represents the reducing agent, the simplest possible reaction representing development is



A suitable reducing agent must not be too weak, nor must it be so strong as to act on the unexposed silver grains. Many of the reducing agents used in photography are structurally related to benzene, C_6H_6 , having one or more of the hydrogens replaced by a hydroxyl, OH, or an amino, NH_2 , group. The location of the substituted groups in the benzene ring is of importance. If they are adjacent, the compound name is prefixed by *ortho*; if next but one, it is prefixed by *meta*; if opposite, it is prefixed by *para*. For example, the familiar developing agent hydroquinone is chemically known as paradihydroxybenzene, and *p*-phenylenediamine is para-diaminobenzene. A developer usually contains one or more of these reducing agents in a water solution with some alkali, usually sodium carbonate, which acts as an accelerator, a bromide as a restrainer, and a sulphite as a preservative. The addition of both accelerator and a bromide restrainer makes the chemical action more independent of the amount of bromide introduced during the developing process. It is most convenient to use the prepared developers marketed by several manufacturers. Eastman's D-72 or Dektol is suitable for positive prints as well as negative emulsions. Developer D-19 is often preferred for spectrograms, as it gives a clearer background and greater contrast. The stock solutions should be kept in a nearly full bottle with a tightly fitting rubber stopper. Under these conditions they will keep for several months. The working solution is prepared by diluting the stock solution according to instructions and should also be kept in a rubber-stoppered bottle at all times when it is not being used. The plate is immersed in the developer, emulsion side up, and the solution agitated occasionally to make the development proceed uniformly. After developing for the specified time, the plate is immersed for about 1 min in a chrome alum solution to harden the gelatin and to stop development. After this it is immersed in a fixing bath, which consists principally of a sodium thiosulphate, "hypo," solution which usually also contains some sodium sulphite, potassium

alum, and citric acid. The thiosulphate reacts with the unreduced silver halides and converts them into soluble salts according to a formula like the following for the bromide:



The fixing should proceed for a few minutes after the plate is cleared of all unreduced silver salts. The washing process which follows removes the various soluble chemicals from the gelatin. At least 20 min in running water is required, a longer time doing no harm. Paper prints require a much longer time, up to 2 hours, in order to remove the salts from the paper base. After washing is completed, the surplus water is sponged off with a viscose sponge, and the plate is placed in a dry, dust-free atmosphere until it is completely dry.

The density of the silver deposit is defined as the common logarithm of the opacity, the latter being the reciprocal of the transmissivity. This is measured by means of a densitometer, which employs a thermopile or a photoelectric cell to determine the intensity of the light transmitted. If i_0 is the intensity transmitted by a clear portion of the plate, and i is the intensity transmitted by the photographic deposit, the density is given by the formula

$$D = \log_{10} \frac{i_0}{i}$$

The relative intensities i_0 and i are generally determined from three deflections of the recording galvanometer, namely, d_0 for the zero or "dark" reading, d_c for the clear plate, and d for the deposit. Then

$$D = \log_{10} \frac{d_0 - d_c}{d_0 - d}$$

The density D varies with the exposure E , which is the product of the intensity I to which the plate was exposed and the time t of exposure. A graph of D as a function of $\log E$, as in Fig. 43, gives the characteristic curve, also known as the Hurter-Driffeld or H-D curve. A graph of D against $\log I$ at constant t , or of D against $\log t$ at constant I , gives a qualitatively similar curve. The portion AC is called the toe of the curve; the portion FG , the shoulder. The approximately straight portion CF between them is the

region of correct exposure. The straight line portion may be represented by the equation

$$D = \gamma(\log E - \log E_B)$$

The quantity E_B is called the **inertia** of the emulsion, while γ is called the **contrast** or simply **gamma**. The value of E_B depends principally on the kind of emulsion employed, whereas gamma depends not only on the emulsion but on the development. It in-

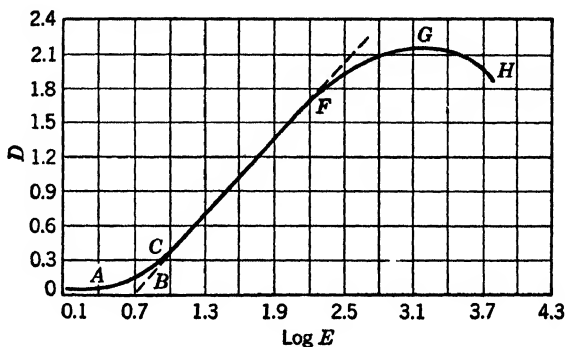


FIG. 43. Characteristic curve of a photographic plate, showing the toe AC , or region of underexposure, the straight line portion CF , the shoulder FG , and the region of solarization GH .

creases with the time of development to a maximum which is characteristic of the emulsion. A gamma of nearly unity is considered desirable for true reproduction of tone values.

A speed index S of the emulsion can be derived from the H-D curve in a variety of ways. The oldest of these, the H-D speed, is defined by $S = 10/E_B$, the inertia E_B being in meter-candle-seconds, mcs. Another speed index, the **Scheiner number**, utilizes the exposure E_A required to produce a minimum perceptible blackening. It is defined by the formula $S' = 10 \log K'/E_A$, the constant K' being different in the several Scheiner systems. The German DIN system is similar to the Scheiner except that, instead of E_A , the exposure E_D required to produce a density of 0.1 above the fog level is employed. Note that, in both the Scheiner and the DIN systems, a decrease of E_A or of E_D to half its value causes the speed index to increase by 3 units, i.e., by $10 \log 2$.

The popular **Weston speed** rating is based on the exposure E_W which is required to produce a density D equal to the gamma obtained with a standardized development. The Weston speed S'' is given by $S'' = 4/E_W$. This speed index has the advantage that the speeds of various emulsions are compared in the range of densities usually employed in practice, instead of in the region of underexposure. There is no accurate correlation between all these speed indices, since the characteristic curves of various emulsions differ considerably; for example, an emulsion having a lower threshold speed or Scheiner rating may have a sufficiently high gamma so as actually to require a shorter exposure to produce a given useful density in the normal range of exposure.

An ordinary plate or film is not sensitive to wave lengths longer than about 5200 Å. This limit may be extended by the use of certain dyes which act as sensitizers. Thus **orthochromatic** plates are sensitized to 5500 Å or even to 6100 Å. They may still be handled in a darkroom under the illumination of a dim red safe-light. Even so, it is best to avoid exposing them to the darkroom light for too long a time, especially at short distances from the light. **Panchromatic** emulsions are sensitized to about 6800 Å, which includes practically all the visible spectrum. They must be handled and processed in complete darkness. Infrared sensitive plates and films are available which are sensitized to about 14,000 Å. They usually have a region of extremely low sensitivity in the green region of the spectrum. Since the eye is very sensitive to these wave lengths, a safe-light transmitting the proper band of green light may be used.

The sensitivity of the silver halides decreases very slowly as the wave lengths become shorter. The gelatin base, however, begins to absorb at about 2300 Å, causing spectrograms to fade out gradually beyond that wave length. This effect may be overcome by using plates containing a minimum of gelatin, such as Schumann plates or Ilford Q plates, or by the use of fluorescent coatings over the emulsion. One may easily prepare the latter by applying a thin film of a paraffin oil such as Nujol with a pad of chamois. However, more uniform results are obtained with commercially coated plates such as the Eastman O series of plates or films. In any case, the fluorescent coating must be removed before development by some solvent such as ethylene chloride or acetone. Trays made of glass or enamelled metal must be used for this purpose.

Short-wave-length x-ray spectra may again be recorded by an ordinary uncoated emulsion. It is preferable that the emulsion be thick to absorb a greater amount of radiation, since it is evident that the energy which is not absorbed in the emulsion cannot produce any photographic effect.

The useful exposure range, or **latitude**, of an emulsion is defined as the ratio of the exposures at the two ends *C* and *F* of the straight line portion of the characteristic curve. In the case of process plates, used principally for copying, this is about 5 to 10, but it may be as large as 250 in the soft portrait emulsions.

In pictorial photography, it is generally desirable to keep the important parts of the picture within the straight line portion of the characteristic. Spectrographic records may, and usually do, run over into the regions of under- and overexposure. However, for the most accurate comparison of relative intensities, one should regulate the exposure so that comparisons are made in the linear region, and one should choose an emulsion having as large a gamma as is consistent with this requirement.

Examination by means of a microscope or magnifier of moderate power shows that all photographic images are more or less grainy. This **graininess** is only partially due to the original size of the silver halide grains, being primarily a result of aggregation or clumping of silver grains during the process of development. Special fine-grain developers such as Eastman Microdol reduce granularity to a marked degree. It seems important to keep the temperature of all baths nearly the same, preferably near 20°C or 68°F. For the best results development should not be carried to a high value of gamma.

The **resolving power** of an emulsion depends partly on graininess and partly on turbidity and contrast. It is usually denoted by the number of lines per millimeter that can just be resolved, using a test object made up of black lines separated by white spaces of equal width. Turbidity is the spreading of an image beyond its normal boundaries because of reflection and diffraction of light by the grains of halide in the emulsion. This leads to a decrease in resolving power. Resolving power varies from about 40 or 50 for high-speed emulsions to about 100 or 150 for the finest-grained plates having a high gamma.

Bunsen and Roscoe in 1862 thought that they had established a reciprocity law indicating that a given density results from a given

total exposure $E = It$ regardless of the values of I and t during the exposure. Thus, in a simple case, an intensity of twice a given value would require just half the exposure time to produce the same density. This is a commonly used rule in photographic practice. At low and high intensities, however, the reciprocity law ceases to hold. This may be shown by curves such as in Fig. 44, which indicate the $\log E$, required to give each of several densities, plotted against the intensity. If the reciprocity law were strictly true,

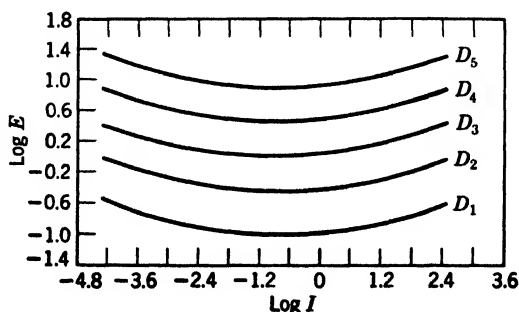


FIG. 44. Failure of the reciprocity law. The exposure E required to obtain any given density, such as D_1 , increases as the intensity I becomes very low and also when it becomes very high.

these curves would all be straight horizontal lines. Instead it is seen that an appreciably greater exposure is required at low or at high intensities. This effect is quite large at low intensities in the case of some emulsions.

Another type of reciprocity law failure applies to intermittent exposures. It is found that intermittent exposures do not produce the same effect as an equivalent continuous exposure extending over the same total time. In most cases, however, one may use an intermittent exposure of high enough frequency as a convenient means for obtaining an intensity calibration of photographic plates. It is always easier to obtain accurately graded times of exposure than to obtain similar steps in intensity. For continuous exposures, the effects are not equivalent in the two cases because of failure of the reciprocity law. However, if a sector disk rotating above the **critical frequency** is used, a variable sector opening, which really varies the *time* of exposure, may be considered as producing equivalent changes in the *intensity* of the light. This has been

verified for many emulsions. Some plates do not seem to have a critical frequency, according to Vincent and Sawyer.

For some purposes, such as quantitative spectrochemical analysis, it is necessary to measure the relative intensities of two or more spectral lines. This may be accomplished with the aid of a D versus $\log I$ calibration curve for the plate used. The lines whose intensities are to be compared should not be too different in wave length, since gamma and sensitivity are not constant throughout the spectrum. If the exposures of both lines are such that their densities are on the straight line portion of the characteristic, then

$$\log \frac{I_1}{I_2} = \frac{D_1 - D_2}{\gamma}$$

and a determination of gamma suffices to convert the density difference into an intensity ratio. The value of gamma may be obtained conversely by the use of two intensities which have a known ratio. If some of the densities are not on the straight line portion, however, the complete characteristic curve must be plotted with the aid of a sufficient number of graduated intensities. Such a calibration of the plate may be simultaneously accomplished with the recording of the spectrogram by the use of either (a) a stepped slit, (b) a logarithmic sector, (c) a stepped sector, (d) a neutral absorbing wedge, or (e) spectral lines in close groups whose relative intensities are known. For experimental details, the reader is referred to Sawyer's *Experimental Spectroscopy*.

APPENDIX

TABLE 1

FUNDAMENTAL PHYSICAL CONSTANTS¹

c	Velocity of light	2.9976×10^{10} cm sec ⁻¹
e	Electronic charge	4.8024×10^{-10} esu
m	Electronic mass	9.1055×10^{-28} gm 0.0005487 amu
e/m	Specific charge of electron	1.75936×10^7 emu gm ⁻¹
M_1	Grams per atomic mass unit	1.660×10^{-24} gm
h	Planck's constant	6.6234×10^{-27} erg sec
N_0	Avogadro's number	6.0235×10^{23} mole ⁻¹
k	Boltzmann's constant	1.3803×10^{-16} erg deg ⁻¹
R	Rydberg's wave number	$109,737.30$ cm ⁻¹
M_H/m	Mass ratio of proton to electron	1836.57
M_{1c^2}	Rest energy per atomic mass unit	931.04 mev
F	Faraday	9652.2 emu equiv ⁻¹
α	Fine structure constant	7.2981×10^{-3}
	Wien's displacement-law constant	0.289715 cm deg
	Stefan's constant	5.6724×10^{-5} erg cm ⁻² deg ⁻⁴ sec ⁻¹
	First radiation constant ($8\pi hc$)	4.9902×10^{-15} erg cm
	Second radiation constant (hc/k)	1.43847 cm deg
	Least mechanical equivalent of light at 0.555 μ	0.00161 watt/lumen
	Ratio grating to Siegbahn scale	1.002030
H	Mass of hydrogen atom	1.008131 amu
D	Mass of deuterium atom	2.014725 amu
He	Mass of helium atom	4.003860 amu

$$\lambda = \frac{12,394.2}{V} \text{ \AA} \quad \text{or} \quad \bar{\nu} = 8068.2 \text{ V cm}^{-1}$$

¹ DUMOND and COHEN, *Reviews of Modern Physics*, **20**, 82-108 (1948).

TABLE 2

SOME USEFUL WAVE LENGTHS

(Wave lengths in air at 15°C and 76 cm Hg barometric pressure, expressed in angstrom units)

<i>Hydrogen</i>	<i>Barium-cored Carbon Arc</i>	<i>Miscellaneous</i>	
6562.8		6707.84	Li
4861.3	6496.90	6438.4696	Cd standard
4340.5	6141.72	6103.64	Li
4101.7	5853.68	5895.923	Na
	5535.55	5889.953	Na
<i>Mercury</i>	4934.09	5535.55	Ba
6907.16	4554.04	5085.82	Cd
6234.37		4799.92	Cd
6123.27	<i>Fraunhofer Lines</i>	4678.16	Cd
6072.64	B 6870.0 O ₂	3968.47	Ca II
5790.65	C 6562.8 H	3933.67	Ca II
5769.59	D ₁ 5895.92 Na		
5460.740	D ₂ 5889.95 Na		
4916.04	E 5270.1 Fe, Ca		
4358.35	F 4861.3 H		
4347.50	G' 4340.5 H		
4339.24	H 3968.5 Ca II		
4077.81			
4046.56			

TABLE 3
SPECTRUM OF THE MERCURY ARC IN THE ULTRAVIOLET

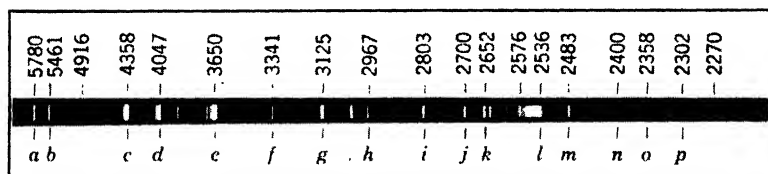


FIG. 45. Mercury-arc spectrogram, showing wave lengths of principal lines. More exact values of wave lengths are listed in the accompanying table.

Lines	Approx.	Wave Lengths of Lines in Group
<i>a</i>	5780	5790.6, 5769.6
<i>b</i>	5461	
<i>c</i>	4358	4358.4, 4347.5, 4339.2
<i>d</i>	4047	
<i>e</i>	3650	3654.8, 3650.1
<i>f</i>	3341	
<i>g</i>	3125	3131.8, 3131.5, 3125.7
<i>h</i>	2967	
<i>i</i>	2803	2804.5, 2803.5
<i>j</i>	2753	2752.8, 2752.2
<i>k</i>	2652	2655.1, 2653.7, 2652.0
<i>l</i>	2537	
<i>m</i>	2483	2484.8, 2483.8, 2482.7, 2482.0
<i>n</i>	2400	2399.7, 2399.4
<i>o</i>	2358	
<i>p</i>	2302	
	2002	The plate must be especially
	1976	sensitized to the ultraviolet in
	1942	order to record the last four
	1850	lines listed here.

TABLE 4

REFLECTIVITY

(All values expressed as percentages)

a. Specular Reflectivity

Wave Length, Microns	Steel	Speculum (Mirror Bronze)	Silver	Aluminum	Glass Mirror (Silver Back)
0.25	38	30	34	85	
0.35	49	51	68	89	
0.45	52	56	90	90	86
0.55	55	64	93	89	88
0.70	58	67	95	86	90
1.00	63	71	97	92	
4.00	88	89	98	91	
14.00	96	94	99	98	

b. Diffuse Reflectivity for White Light

Substance	Diffuse Incident Light	Normal Incidence
Magnesium oxide	96	92
Magnesium carbonate	98	93
Plaster of Paris	91	87
Matt white Celluloid	80-85	75-80
White blotting paper	80-85	75-80

TABLE 5

INDICES OF REFRACTION

a. Gases at Normal Temperature and Pressure and 5893 Å

Air	1.0002926
Hydrogen	1.000132
Nitrogen	1.000296
Oxygen	1.000271
Carbon dioxide	1.000448
Methane	1.000444

b. Liquids at 20°C and 5893 Å (Indices relative to Air)

Water	1.33299 ¹
Alcohol	1.36048
Benzene	1.50142
Carbon disulphide	1.6295
Monobromonaphthalin	1.658
Canada balsam	1.530
Ethyl ether	1.3497

¹ 1.33192 at 30°C.

SOLIDS AT 5893 Å

c. Miscellaneous Anisotropic Solids

Fused quartz	1.45843
Gold	0.42
Platinum	2.04
Silver	0.18
Diamond	2.4173
Glass	
Ordinary crown	1.5159
Light flint	1.5710
Dense flint	1.7541
Densest flint	1.9626

d. Optical Glass

Wave Length, Å	Ordinary Crown	Borosilicate Crown	Barium Flint	Light Flint	Me- dium Flint
Hg 4046.8	1.5319	1.5382	1.5885	1.6051	1.6579
H 4340.7	1.5282	1.5347	1.5833	1.5986	1.6497
Hg 4358.6	1.5280	1.5345	1.5830	1.5983	1.6493
H 4861.5	1.5233	1.5301	1.5765	1.5903	1.6394
Hg 4916.4	1.5228	1.5297	1.5759	1.5896	1.6385
Hg 5461.0	1.5193	1.5263	1.5711	1.5838	1.6314
Hg 5790.5	1.5176	1.5248	1.5688	1.5811	1.6282
Na 5893.2	1.5171	1.5243	1.5682	1.5804	1.6273
H 6563.0	1.5146	1.5219	1.5648	1.5764	1.6224
K 7682	1.5116	1.5191	1.5610	1.5718	1.6170

e. Crystals

Uniaxial

	n_o	n_e
Calcite	1.6584	1.4864
Quartz	1.5442	1.5533
Ice	1.3091	1.3104
Sodium nitrate	1.5874	1.3361

Biaxial

	α	β	γ
Mica, muscovite	1.5601	1.5936	1.5977
Potassium nitrate	1.3346	1.5056	1.5061
Sucrose	1.5397	1.5667	1.5716
Tartaric acid	1.4953	1.5353	1.6046
Rochelle salt	1.48984	1.49166	1.49525

TABLE 6
STANDARD VISIBILITY FUNCTION ¹
(Relative to $V_\lambda = 1.000$ at $\lambda 0.555 \mu$)

Wave Length, Micron	V_λ	Wave Length, Micron	V_λ
0.38	40×10^{-6}	0.58	0.8700
0.40	40×10^{-5}	0.60	0.6310
0.42	40×10^{-4}	0.62	0.3810
0.44	0.0230	0.64	0.1750
0.46	0.0600	0.66	0.0610
0.48	0.1390	0.68	0.0170
0.50	0.3230	0.70	41×10^{-4}
0.52	0.7100	0.72	105×10^{-5}
0.54	0.9540	0.74	250×10^{-6}
0.56	0.9950	0.76	60×10^{-6}

¹ JUDD, *Bureau of Standards Journal of Research*, **6**, 465 (1931).

CORRECTED ANSWERS

ANSWERS TO PROBLEMS

CHAPTER 1, PAGE 9

1. (a) $40^\circ 30'$; (b) $19^\circ 30'$. 2. $-2.85, 1.63, 0$. 3. (a) $19^\circ 35'$; (b) 6.68 ft.
4. 0.554.

CHAPTER 2, PAGE 18

1. (b) $(1 + 1/M)f$. 2. x_0 and x_0' are, respectively, the distances from the principal foci F and F' to the conjugate points from which p and p' are measured. 4. (b) $\overline{XN}:\overline{XN'} = s:s'$. 5. $f = f' = 26.67$ cm; $p = 6.67$ cm; $p' = -13.33$ cm.

CHAPTER 3, PAGE 33

1.

Fraunhofer Line		$h = 0$	$h = 10$ mm	$h = 15$ mm
C	u_2' s_2' 96.78 mm	$5^\circ 56' 30''$ 94.44 mm	$9^\circ 6' 35''$ 91.44 mm
D	u_2' s_2' 96.33	$5^\circ 58' 21''$ 94.00	$9^\circ 9' 13''$ 91.01
F	u_2' s_2' 95.18	$6^\circ 2' 54''$ 92.85	$9^\circ 16' 8''$ 89.83
G'	u_2' s_2' 94.29	$6^\circ 6' 16''$ 91.98	$9^\circ 21' 33''$ 88.94

2. (a) 12 cm; (b) 20 cm. 3. -20 cm. 4. 83.3 cm. 5. $f = 80.9$ cm; $p = 1.214$ cm; $p' = -2.43$ cm. 6. 84 cm; 4.2 cm. 7. -77.4 mm. 8. -61.6 mm.
9. -100 cm; 3. 10. -104.0 cm; 3.14. 11. $f = 36.89$ cm; $p = 15.22$ cm; $p' = -25.40$ cm. 12. (a) $p = -2.21$ mm; $p' = -1.11$ mm; $f = 97.43$ mm; (b) 97.43 mm; (c) 10.26 D; 10.38 D; (d) $s_2' = 186.92$ mm; $\beta = 0.9303$; $s' = 188.03$ mm. 13. $s_2' = 52.78$ cm; 13.89 cm. 14. 2 mi. 15. 136.9 cm. 16. $4f + PP'$; interchange values of s and s' . 17. $s_2' = 14.29$ cm; real, inverted; 4.29 cm. 18. $p = 4$ cm; $p' = -8$ cm; $f = f' = 12$ cm. 19. $s_2' = 7.77$ cm; $\beta = 0.530$. 20. $f = 10.00$ cm; $p = 10$ cm; $p' = -6.25$ cm. 21. $s_2' = 580$ cm; 150 cm. 22. (a) $p = 1.836$ mm; $p' = -4.40$ mm; $f = 139.7$ mm; (b) $+7.4$ D;

- (c) 43.6 mm; -104.6 mm. **23.** $\frac{n_1 r^2}{(n_1 - 1)(n_2 - n_1)l_1 + n_1(2n_1 - n_2 - 1)r}$. **24.** 5.08 mm; 50.8 mm. **25.** $p = 5.75$ cm; $p' = -7.58$ cm; $f = 27.9$ cm.

CHAPTER 4, PAGE 51

- 1.** (L.S.A.)_D = 5.32 mm; sine cond. = 3.15 mm; coma = -2.17 mm; chrom. F-C 1.60 mm. **2.** (a) -0.697; (b) 0.613 cm. **3.** $R_1 = 7.00$ cm; $R_2 = -5.25$ cm. **4.** (a) $h^2 n^2 / 2f(n-1)^2$; (b) $h^2(n^3 - 2n^2 + 2) / 2nf(n-1)^2$. **5.** (a) $s_1' = 43.3$ cm; $s_2' = 208.3$ cm; (b) $l = 3.96$ cm. **6.** 7.85 cm; -12.81 cm. **7.** $\beta = f/x$; object must be distant enough so that chromatic displacements in principal focus F are negligible.

CHAPTER 5, PAGE 59

- 1.** 2 ft 11 in.; bottom edge 2 ft 8 in. above the floor. **2.** 30 mm. **3.** 6.67; $f/3.33$. **4.** (a) The rim of the first lens; 1.50-cm radius; 12 cm to left of second lens; 2.25-cm radius; (b) $f/3.43$; (c) 24 cm to right of first lens; 3.0-cm radius; second lens; 1-cm radius. **5.** Entrance pupil at objective; exit pupil 3.2 cm to left of eyepiece, 0.5-cm radius; $7^\circ 9'$; $34^\circ 42'$. **6.** (a) Objective, 20-mm diam.; (b) 16.25 mm to right of eyepiece, 1.67-mm diam.; (c) diaphragm, 7-mm diam.; (d) $2^\circ 13'$. **8.** (a) 4.77; (b) 2.44; because this varies somewhat with location of object. **9.** 4.35 meters to 9.70 meters.

CHAPTER 6, PAGE 72

- 1.** 28.8 ft-c. **2.** 100 ft-c. **3.** (a) 22.2 ft-c; (b) 0.309 lumen. **4.** 6.29 ft-c. **5.** (a) 17.48 millilamberts; (b) 5.57×10^{-3} cp/cm.² **6.** 64 sec. **7.** 0.65 ft-c. **8.** 162 ft-c. **9.** (a) 70.7%; (b) 67.9%.

CHAPTER 7, PAGE 98

- 1.** (a) -1.71 D; (b) 18.9 cm from eye. **2.** 15.3 cm to 755 cm. **3.** +3.28 D. **4.** +10.43 D. **5.** 200 cm. **6.** 1.12 ft. **7.** 0.8 cm from first lens. **8.** -1.5 cm. **9.** (a) 6 cm; (b) 30 cm. **10.** 2.67 mm. **11.** (a) 12.5; (b) 53.33; (c) 667. **12.** 11.96 cm. **13.** 10.67 cm; -2.67 cm. **14.** (a) 6.67; (b) 1.43 in. **15.** 5.55 cm. **16.** 1.539. **18.** $74^\circ 36' 42''$. **19.** (a) $3^\circ 11'$; (b) 5.56 cm. **20.** $3^\circ 46'$. **21.** $5^\circ 7'$; $1^\circ 58'$. **22.** $60^\circ 24'$. **23.** $R > 10^4$; recip. disp. = 12.5 Å/mm; desirable $R \cong 20,000$; recip. disp. $\cong 6$ Å/mm. **24.** (a) 31.3; (b) 2.71 mm beyond eye lens, 0.8-mm diam.; (c) $2^\circ 16'$; (d) $5.53''$.

CHAPTER 8, PAGE 105

- 1.** 332.5 cm. **2.** 3.33 cm; 0.6; 0.4. **3.** (a) 447 meters; 4530 meters. **4.** Right eye: +7.87 D.S., $\Delta 5.9$ base out; left eye: same.

CHAPTER 9, PAGE 114

1. $42^h 28^m 57.2^s$. 2. 23.4 km. 3. See Anderson, *Journal of the Optical Society of America*, **31**, 187 (1941). 4. 1.9760×10^{10} cm/sec; 1.9418×10^{10} cm/sec; 1.8421×10^{10} cm/sec; 1.7841×10^{10} cm/sec.

CHAPTER 10, PAGE 154

1. 13.99 statvolts/cm. 2. 5.893×10^{-4} cm. 3. 41.0×10^{-4} cm. 4. 2.36×10^{-4} cm. 5. 76.5 cm. 6. 55.2 cm. 7. (a) Max. int. when $2t \cos \phi = (m + \frac{1}{2})\lambda$; min. when $2t \cos \phi = m\lambda$; interchange max. and min. if $n_1 < n_2 < n_3$; (b) at thin wedge; at infinity; (c) selective interference (channelled spectrum with few fringes). 8. In microns: 0.667; 0.572; 0.500; 0.445; 0.400. 9. 0.107μ ; dark purple; more visual reflectance and bluer color. 10. 127.3. 11. In microns: 0.572; 0.545; 0.522; 0.500. 12. 18.5μ . 14. $(\sin \pi Dw) / (\pi Dw)$; see Michelson, *Studies in Optics*, page 37. 15. (a) $e^{-(\pi Dw)^2}$; (b) 0.088 cm^{-1} ; 0.022 \AA . 16. 763,000.

CHAPTER 11, PAGE 191

1. (a) 0.091 cm; 1.34; 0.139 cm; 0.78; (b) 12.25 cm; 1.34; 18.71 cm; 0.78. 3. *Fraunhofer pattern*: same distribution law but narrower maxima as the slit is widened. *Fresnel pattern*: change in distribution from *diffusion*, through *outer* bands, to *inner* bands. The extent of the pattern narrows, then widens, as the slit width increases. 4. 24.4 cm. 5. 3.86 in. 6. 4.88×10^{-4} -in. diam. 7. (a) 1.74 in.; 21.2. 8. 488-cm diam. 9. $1.05'$. 10. $f/32.8$. 11. Take sum of squares of x and y components of chords of vibration circle. 12. (a) 0.516 in.; (b) 35.6. 13. 32.6 cm. 14. $R > 235,720$; $D > 1.165 \times 10^6$ rad/cm without magnification. 15. (a) 4.86 cm; (b) 11.11 meters. 16. $14^\circ 44'$; $30^\circ 33'$. 17. $19^\circ 57'$; $28^\circ 51'$; $17^\circ 11'$.

CHAPTER 12, PAGE 225

1. Ordinary ray; $i > 67^\circ 21'$. 2. (a) 1.732:1; (b) 3:1. 3. $l = m \cdot 0.06548$ mm. 4. Wave lengths of minimum intensity equal to optical path difference divided by successive integers; (b) path difference is $(\mu_2 - \mu_1)l$ as compared with $2\mu t$. 5. (a) 0.688; 0.492; 0.383μ ; (b) 0.573; 0.430μ . 6. (a) 0.00023 mm; (b) $150^\circ 33'$. 7. $5^\circ 48' 48''$. 8. (a) 0.10638; (b) $29^\circ 37' 10''$. Derive formula or see Schuster, *The Theory of Optics*. 9. $11^\circ 37' 36''$. 10 and 11. See Drude, *Theory of Optics*.

CHAPTER 13, PAGE 231

1. 75.33 gm./lit. 2. 5.66×10^{-7} . 3. (a) 4.14 mm; (b) 64.1° . 4. 98.82%. 5. 1.0038 mm; borosilicate crown; 1.0199 mm.

CHAPTER 14, PAGE 241

1. (a) 0.158; 0.152; (b) 1.000317. 2. 1.000431; 85×10^{-6} low. 3. 1.84×10^8 as compared with 2.636×10^8 . 4. (a) Page 190; (b) 3.65×10^{-6} . 5. (a) $6.24'$; (b) 0.906 cm.

CHAPTER 15, PAGE 257

1. 77.8° . 2. 154° . 3. $46^\circ 59' 20''$ or $55^\circ 52' 30''$. 4. $76^\circ 49'$; $26^\circ 50'$.
 5. (a) $157^\circ 37'$; (b) $(p + 0.938)\lambda/2$, $p = 0, 1, 2, 3, \dots$.

CHAPTER 16, PAGE 269

1. $T < 2741^\circ\text{K}$. 2. $u_\lambda = C_1' T^{1/2} \lambda^{-5.5} (e^{hc/\lambda kT} - 1)^{-1}$; relatively higher at shorter wave lengths. 3. 1900°K . 4. (a) 375.1 erg/cm^2 ; (b) $8.956 \times 10^{11} \text{ ergs/cm}^2 \text{ sec}$. 5. $6.695 \times 10^{-27} \text{ erg sec}$; $1.395 \times 10^{-16} \text{ erg/deg K}$.

CHAPTER 17, PAGE 298

1. Lyman: $82,258 \text{ cm}^{-1}$; 1215.6 \AA ; $109,677 \text{ cm}^{-1}$; 911.7 \AA . Balmer: $15,234 \text{ cm}^{-1}$; 6564.1 \AA ; $27,419 \text{ cm}^{-1}$; 3646.7 \AA . Paschen: 5332 cm^{-1} ; $18,754 \text{ \AA}$; $12,187 \text{ cm}^{-1}$; 8205.1 \AA . 2. (a) 4.14 cm^{-1} ; (b) 1.784 \AA . 3. $1.76034 \times 10^7 \text{ emu/gm}$. 4. $40,468 \text{ cm}^{-1}$; 4.77469 ; $40,138.3 \text{ cm}^{-1}$.

CHAPTER 18, PAGE 311

1. 1.422 \AA . 2. (b) $\Delta \left(\frac{\nu}{R} \right) = \frac{\alpha^2(Z - \sigma_2)^4}{n^3 l(l+1)}$. 3. (a) 30.99 XU ; (b) constant = $\frac{3hc}{e} \times 10^{10} \text{ XU/kv}$. 4. 24.5% . 5. $8.08 \times 10^{-8} \text{ erg}$; $9.34 \times 10^{-8} \text{ erg}$.

CHAPTER 19, PAGE 324

1. (p) 0, $\pm 1/6$; (s) 1, $\pm 7/6$, $\pm 4/3$ for $^1P_1 - ^3D_2$. 2. $78,000 \text{ volts/cm}$.
 3. $22,200 \text{ gauss}$.

CHAPTER 20, PAGE 330

1. 0.143; 0.256; 0.083. $x = 0.298$; $y = 0.530$; $z = 0.172$. 2. $511.0 \text{ m}\mu$; 34.9% ; 0.256. 3. 0.041; 0.605; 0.354. 4. See Hardy, *Handbook of Colorimetry*

BIBLIOGRAPHY

a. GEOMETRICAL OPTICS

- CONRADY, *Applied Optics*, Oxford Press, 1929.
DRUDE, *Theory of Optics*, Longmans, Green, 1920.
GLEICHEN, *The Theory of Modern Optical Instruments*, H. M. Stationery Office, 1921.
HARDY and PERRIN, *The Principles of Optics*, McGraw-Hill, 1932.
JACOBS, *Fundamentals of Optical Engineering*, McGraw-Hill, 1943.
LUMMER, *Contributions to Photographic Optics*, Macmillan, 1900.
MOFFIT, Telescope Objective Design, *Journal of the Optical Society of America*, **11**, 147 (1925).
MONK, *Light, Principles and Experiments*, McGraw-Hill, 1937.
MOON, *Scientific Basis of Illuminating Engineering*, McGraw-Hill, 1936.
NUTTING, *Outlines of Applied Optics*, Blakiston, 1912.
SOUTHALL, *Mirrors, Prisms, and Lenses*, Macmillan, 1933.
SOUTHALL, *The Principles and Methods of Geometrical Optics*, Macmillan, 1910.
STEINHEIL and VOIT, *Applied Optics*, Blackie, 1918.
TAYLOR, *A System of Applied Optics*, Macmillan, 1906.
U. S. BUREAU OF STANDARDS, *Scientific Papers*, Numbers 461 and 550.
VON ROHR, *Geometrical Investigation of the Formation of Images in Optical Instruments*, H. M. Stationery Office, 1920.
WALSH, *Photometry*, Constable, 1926.
WHITTAKER, *The Theory of Optical Instruments*, Cambridge Press, 1915.

b. PHYSICAL OPTICS

- BORN, *Optik*, Julius Springer, 1933.
BRUHAT, *Cours d'Optique*, Masson & Cie, 1931.
DRUDE, *Theory of Optics*, Longmans, Green, 1920.
EDSER, *Light for Students*, Macmillan, 1946.
FOERSTERLING, *Lehrbuch der Optik*, S. Hirzel, 1928.
GEHRCKE, *Handbuch der Physikalischen Optik*, J. A. Barth, 1927.
GRIMSEHL, *Optics*, Blackie, 1933.
HOUSTOUN, *A Treatise on Light*, Longmans, Green, 1945.
JENKINS and WHITE, *Fundamentals of Physical Optics*, McGraw-Hill, 1937.
MEYER, *Diffraction of Light, X-rays, and Material Particles*, University of Chicago Press, 1934.
MICHELSON, *Light Waves and Their Uses*, University of Chicago Press, 1903.
MICHELSON, *Studies in Optics*, University of Chicago Press, 1927.
MONK, *Light, Principles and Experiments*, McGraw-Hill, 1937.
PRESTON, *Light*, Macmillan, 1928.
ROBERTSON, *Introduction to Physical Optics*, D. Van Nostrand, 1929.
SCHUSTER, *Introduction to the Theory of Optics*, E. Arnold, 1904.

WIEN-HARMS, *Handbuch der Experimental Physik*, Vol. 18, Akademische Verlagsgesellschaft, 1928.

WOOD, *Physical Optics*, Macmillan, 1934.

c. RADIATION AND SPECTRA

CLARK, *Applied X-rays*, McGraw-Hill, 1940.

DRUDE, *Theory of Optics*, Longmans, Green, 1920.

FOERSTERLING, *Lehrbuch der Optik*, S. Hirzel, 1928.

FORSYTHE, *Measurement of Radiant Energy*, McGraw-Hill, 1937.

GEHRCKE, *Handbuch der Physikalischen Optik*, J. A. Barth, 1927.

HARDY, *Handbook of Colorimetry*, Technology Press, 1936.

HERZBERG, *Atomic Spectra and Atomic Structure*, Prentice-Hall, 1937.

MOON, *Scientific Basis of Illuminating Engineering*, McGraw-Hill, 1936.

RICHTMYER and KENNARD, *Introduction to Modern Physics*, 4th ed., McGraw-Hill, 1947.

SAWYER, *Experimental Spectroscopy*, Prentice-Hall, 1944.

TOLANSKY, *High Resolution Spectroscopy*, Pitman, 1947.

WHITE, *Introduction to Atomic Spectra*, McGraw-Hill, 1934.

WOOD, *Physical Optics*, Macmillan, 1934.

d. EXPERIMENTAL OPTICS

BRODE, *Chemical Spectroscopy*, John Wiley, 1943.

CLARK, *Applied X-rays*, McGraw-Hill, 1940.

GERLACH and SCHWEITZER, *Chemical Analysis by Means of the Emission Spectrum*, Hilger, 1929.

HARNWELL, Source for the Balmer Series, *American Physics Teacher*, **3**, 185 (1935).

HENNEY and DUDLEY, *Handbook of Photography*, Whittlesey House, 1939.

HILGER, *Iron Spectrum* (charts), *Hilger E. 1. Spectrograph*, Adam Hilger, 1928.

M.I.T. *Tables of Wave Lengths*, Technology Press, John Wiley, 1939.

MILLER, *Laboratory Physics*, Ginn, 1932.

MONK, *Light, Principles and Experiments*, McGraw-Hill, 1937.

PFUND, Interference Scheme for Measuring Cell Depth, *Journal of the Optical Society of America*, **23**, 416 (1933).

PIERCE, TORRES, and MARSHALL, Identification of Lines in Qualitative Spectrochemical Analysis, *Industrial Engineering Chemistry*, **2**, 191 (1939).

SAWYER, *Experimental Spectroscopy*, Prentice-Hall, 1944.

SIEGBAHN, *Spectroscopy of X-rays*, Julius Springer, 1931.

STRONG, *Procedures in Experimental Physics*, Prentice-Hall, 1939.

TAYLOR, *College Manual of Optics*, Ginn, 1924.

WAGNER, *Experimental Optics*, John Wiley, 1945.

WATSON, *Practical Physics*, Longmans, Green, 1917.

e. LABORATORY ARTS

CEMENTING OF OPTICAL PARTS

HARDY and PERRIN, *Principles of Optics*, McGraw-Hill, 1937, page 346.

Ordnance Document 2037, Government Printing Office, 1924, pages 255-6.

CROSS-HAIRS, MOUNTING OF

MONK, *Light, Principles and Experiments*, McGraw-Hill, 1937, page 435.

ELECTRON MULTIPLIER TUBES

ALLEN, *Physical Review*, **55**, 966-71 (1939).

WAGNER and FERRIS, *Proceedings of the Institute of Radio Engineers*, **29**, 598 (1941).

FILTERS

CORNING GLASS WORKS, *Glass Color Filters*.

EASTMAN KODAK COMPANY, *Written Light Filters*.

STRONG, *Procedures in Experimental Physics*, Prentice-Hall, 1939, pages 341-395.

GEIGER COUNTERS

STRONG, *Procedures in Experimental Physics*, Prentice-Hall, 1939, pages 259-304.

GRINDING AND POLISHING

STRONG, *Procedures in Experimental Physics*, Prentice-Hall, 1939, pages 29-92.

MATERIALS, OPTICAL

HARDY and PERRIN, *Principles of Optics*, McGraw-Hill, 1937, pages 320-335, 384-408.

SAWYER, *Experimental Spectroscopy*, Prentice-Hall, 1944, pages 63-70.

STRONG, *Procedures in Experimental Physics*, Prentice-Hall, 1939, pages 341-396.

MIRRORS, COATING OF

By evaporation

MONK, *Light, Principles and Experiments*, McGraw-Hill, 1937, pages 431-434.

STRONG, *Procedures in Experimental Physics*, Prentice-Hall, 1939, pages 151-187.

By cathode sputtering

MONK, *Light, Principles and Experiments*, McGraw-Hill, 1937, page 434.

STRONG, *Procedures in Experimental Physics*, Prentice-Hall, 1939, pages 151-187.

TAYLOR, *College Manual of Optics*, Ginn, 1924, page 205.

By chemical deposition of silver

MILLER, *Laboratory Physics*, Ginn, 1932, pages 266-270.

TAYLOR, *College Manual of Optics*, Ginn, 1924, pages 201-207.

WOOD, *Physical Optics*, 2nd ed., Macmillan, 1914, page 281.

MISCELLANEOUS PROCESSES IN MANUFACTURE AND ASSEMBLY

- HARDY and PERRIN, *Principles of Optics*, McGraw-Hill, 1937.
JACOBS, *Fundamentals of Optical Engineering*, McGraw-Hill, 1943.
Ordnance Document 2037, Government Printing Office, 1924.
STRONG, *Procedures in Experimental Physics*, Prentice-Hall, 1939.

PHOTOELECTRIC CELLS

- FORSYTHE, *Measurement of Radiant Energy*, McGraw-Hill, 1937, pages 216–245.
HUGHES and DUBRIDGE, *Photoelectric Phenomena*, McGraw-Hill, 1932, Chapters 12–13.
STRONG, *Procedures in Experimental Physics*, Prentice-Hall, 1939, pages 396–448.

PHOTOGRAPHIC TECHNIQUE

- HENNEY and DUDLEY, *Handbook of Photography*, Whittlesey House, 1939.
MACK and MARTIN, *The Photographic Process*, McGraw-Hill, 1939.
MILLER, *Principles of Photographic Reproduction*, Macmillan, 1942.
STRONG, *Procedures in Experimental Physics*, Prentice-Hall, 1939, pages 449–492.

RADIOMETERS

- FORSYTHE, *Measurement of Radiant Energy*, McGraw-Hill, 1937, pages 189–282.
STRONG, *Procedures in Experimental Physics*, Prentice-Hall, 1939, pages 305–340.

SOURCES OF LIGHT

- HARDY and PERRIN, *Principles of Optics*, McGraw-Hill, 1937, pages 160–184.
HARNWELL, *American Physics Teacher*, **3**, 185 (1935).
MOON, *Scientific Basis of Illuminating Engineering*, McGraw-Hill, 1936, pages 73–179.
SAWYER, *Experimental Spectroscopy*, Prentice-Hall, 1944, pages 18–27.
STRONG, *Procedures in Experimental Physics*, Prentice-Hall, 1939, pages 341–358.
TOLANSKY, *High Resolution Spectroscopy*, Pitman, 1947, pages 19–70.
UREY *et al.*, *Review of Scientific Instruments*, **3**, 497 (1932).

SOURCES FOR COLORIMETRY

- HARDY, *Handbook of Colorimetry*, Technology Press, 1936.
MONK, *Light, Principles and Experiments*, McGraw-Hill, 1937, pages 436–438.

VACUUM PUMPS AND GAUGES

- INANANANDA, *High Vacua*, D. Van Nostrand, 1947.
STRONG, *Procedures in Experimental Physics*, Prentice-Hall, 1939, pages 93–151.
SULLIVAN, *Review of Scientific Instruments*, **19**, 1–16 (1948).

INDEX

A

Abbe, numerical aperture, 66
sine condition, 36
Aberrations, 38
Absorption, bands in liquids and solids, 294
Bouguer's law, 8
cell-thickness measurement, 389
Lambert's law, 8
relation to dispersion, 237
Absorption coefficient, 9, 245
Absorptivity, 380
Accommodation, 74
amplitude of, 75
Achromatic doublet, 50
Allen, electron multiplier tubes, 443
Aluminum, for mirrors, 8, 434
interferometer, 152
table of reflectance, 434
Ametropia, 75
Anderson, speed of light, 112
Andrews, diffraction experiments, 185
Angle of projection, 53
Angular aperture, 53
Angular magnification, 16
Angular momentum, Bohr unit, 273, 276
Angular scale collimator, 342
Aniseikonia, 81
Anisometropia, 81
Anisotropic media, 109, 203, 210, 218
Apertures of optical systems, 53
Aperture stop, 53
Aplanatic surface, 5
Arkadiev, diffraction photographs, 156
Astigmatic difference, 47
Astigmatism, 45
measurement of, 343

Astigmatism, of eye, 78
of thin lens, 47
Atomic structure, Bohr theory, 272
characteristic x-rays, 299
line spectra, 271
Axes, of elliptically polarized light, 219, 406
optic, of crystals, 203
orientation of, 407
Axial ratios, 188

B

Babinet compensator, 220, 408
Bacher and Goudsmit, atomic energy states, 279
Bäcklin, x-ray spectra, 310
Band spectra, 270, 286
P, *Q*, and *R* branches, 292
Bartholinus, double refraction, 202
Beam of light, 1
Beams, electro-optical shutter, 320
Beer's law, 348
Bending of lens, 43
Bessel functions, 162
Betelgeuse, diameter of, 123
Biaxial crystals, 203
Bibliography, 441-444
Binocular vision, 101
Binormals, 207
Biquartz, 231, 351
Birefringence, 217
Birge, physical constants, 431
Black-body radiation, 259
Bohr, theory of atomic spectra, 272
Bouguer's law, 82
Bradley, aberration of light, 111
Bragg, correction for refraction, 190
diffraction formula, 189, 409
Brewster angle, 201, 253

Brewster's law, 201
 Brightness, of images, 64
 of surfaces, 63
 Brode, iron-spectrum charts, 396
 Bruhat, interferometry, 153
 Bunsen-Kirchhoff law, 270
 Bunsen-Roscoe reciprocity law, 426

C

Cadmium standard line, 146
 Canal rays, 319
 Candle power, 62
 measurement of, 344
 Cardinal points, 17
 experimental location of, 338
 of combinations, 24
 Cassegrainian telescope, 90
 Cauchy formula, 93, 356
 Channelled spectrum, 127
 experiment, 338
 Chemical effects on x-ray emission,
 310
 Chief rays, 55
 Chromatic aberration, 49
 measurement of, 341
 Chromaticity, 327
 Circle of confusion, 56
 Circular aperture diffraction, 161, 182
 Circularly polarized light, 218
 interference in, 221
 Coblentz galvanometer, 380
 Coddington, factors, 39
 formula for astigmatism, 46
 Coherent waves, 117
 Collinear transformation, 11
 Collision broadening, 145
 Color, measurement of, 325
 Colorimetry, 348
 Colors of thin films, 127
 Color temperature, 268
 Color triangle, 328
 Coma, 41
 Combinations of optical systems, 24
 Compound microscope, 85
 Compton and Allison, x-rays, 301, 417
 Computation of focal length, 31
 Concave grating, Eagle mounting, 385

Concave grating, experiment, 383
 measurement of wave lengths, 386
 Paschen mounting, 385
 Rowland mounting, 385
 Concurrent systems, 18
 Conical refraction, 211
 Conjugate foci, 10, 15
 Conrady, trigonometric ray tracing, 21
 Constant deviation prism, 94
 Contracurrent systems, 18
 Cornu-Jellet prism, 230
 Cornu spiral, 177
 Critical angle, 254, 256
 experiment, 362
 Crystals, lattice constant, 410
 optics of, 203
 structure of, 203, 417
 body-centered cubic, 420
 face-centered cubic, 420
 families of planes, 186
 powder method, 417
 systems, 203
 Curvature of image, 49
 Cylinders, 78

D

DeBroglie wave length, 279
 Debye-Sherrer-Hull method, 417
 Densitometer, 399
 Density, of absorber, 380
 photographic, 400
 Depolarization factor, 334
 Depth, of field, 56
 of focus, 56
 Deuterium, atomic weight of, 374
 dissociation of, by light, 324
 Deviation by prism, 92
 Diaphragms, 53
 Dichroic media, 104, 212
 Dielectric displacement, 107, 194
 Dielectric polarization, 232
 Diffraction, 156-192
 circular aperture, 161
 straight edge, 180
 Diffraction grating, 168, 171
 Diffuse reflection, 8, 434
 Diopeters, 29

Discontinuous spectra, 270
 Dispersion, 93
 electron theory, 235
 experiment, 356
 of grating, 172
 of prism spectroscope, 95
 quantum theory of, 238
 x-ray, 239
 Dispersive power, 50
 Distinctness of interference fringes, 129, 140
 Distortion, 43
 Distribution, spectral, 260
 Dominant wave length, 329
 Doppler-Fizeau broadening, 145
 Doppler-Fizeau effect, 270
 Dorsey, speed of light, 112
 Double refraction, 210
 Double slit, diffraction by, 164
 Young's experiment, 115
 Doublet, achromatic, 50
 Doubly refracting media, 202
 Drude, ray velocity surface, 208
 Duboscq colorimeter, 348
 Duffendack, arc source, 285

E

Einstein, photoelectric equation, 322
 theory of relativity, 140
 Electrical conductivity, 242, 247, 249
 Electromagnetic theory, 106, 194
 reflection and refraction, 199
 speed of waves, 196
 Electronic bands, 291
 Electron microscope, 87
 Electron theory of dispersion, 235
 Elliptically polarized light, 218
 by oblique metallic reflection, 252
 experimental analysis of, 406
 sense of rotation, 409
 Ellipticity at Brewster angle, 254
 Emission, line spectra, 270
 thermal radiation, 259
 ultraviolet, 391
 x-ray, 299
 Emissivity, 259
 Empty magnification, 87

Enantiomorphic structures, 230
 Energy density, 260
 Energy distribution curve, 265
 determination of, 382
 Energy-level diagrams, 277
 mercury, 278
 sodium, 277
 x-ray, 302
 Entrance port, 56
 Entrance pupil, 53
 Ether drift, 138
 Eye, structure and cardinal points of, 73
 Eyepieces, 83
 Exit port, 56
 Exit pupil, 54
 Experiments in optics, 338-421
 bibliography, 441-444
 Extraordinary ray, 212

F

Fabry-Perot interferometer, 147
 adjustment of, 373
 experiment, 373
 wave-length ratio by, 152, 374
 Faraday effect, 312
 Far point, 75
 Farwell, aplanatic surfaces, 5
 Fermat's principle, 3
 Field of view, 55
 Films, interference in, 124
 Fine structure, in line spectra, 276
 in x-ray spectra, 303
 First order theory, 22
 Fizeau, speed of light, 113
 Floating spot, 104
 Flux into optical system, 64
f-number, 54, 67
 Focal length, 10, 15
 combination of systems, 24
 measurement, diverging lens, 339
 thick lens, 342
 ratio, 24
 thick-lens formulas, 28
 Focus, 10
 Foot-candle, 63

Forsythe, measurement of radiant energy, 442

Foucault, rotating mirror, 111

Fractional excess, 152

Fraunhofer diffraction phenomena, 157

Fraunhofer lines in solar spectrum, 281

wave lengths of, 432

Frequency, 108

Fresnel, crystal optics, 207

formulas, 201

reflection formulas, 124, 201

experimental test, 404

Fresnel biprism, 119

Fresnel diffraction phenomena, 157, 172

Fresnel integrals, 176

Fresnel mirrors, 119

Fresnel prism, 228

Fresnel rhomb, 257

Fresnel zones, 182

Fringes, of constant inclination, 127

of constant thickness, 128

Fuessner spark source, 286

Fundamental physical constants, 431

G

Galilean telescope, 88

Gas interferometer, 123

Gauss eyepiece, 358

Gauss points, 17

General x-radiation, 226

Geneva lens gauge, 30

Geometrical optics, bibliography, 441

postulates, 1, 3

g-factor, 318

Ghosts in grating spectrum, 171

Glancing angle, 189

Glass, optical constants, 435

Gold, optical constants, 245

Gratings, reflection, 364

transmission, 168, 364

Gray body, 267

Grimaldi, diffraction, 156, 173

Group velocity, 109

Gurney, latent image, 422

H

Haidinger's fringes, 127

Half-shade angle, 350

Half-shades for polarimetry, 230

Half width of spectral lines, 142

Hardy, color measurement, 325

sources for colorimetry, 444

Hardy and Perrin, applied optics, 441

Harnwell, hydrogen light source, 376

electrodeless discharge, 444

Hartley, persistent lines, 281

Hartmann, dispersion formula, 93

interpolation formula, 396

Henney and Dudley, photography, 444

Herapathite, 212

Herzberg, width of spectral lines, 145

Herschelian telescope, 90

Hilger charts of the iron spectrum, 396

Hönl, x-ray dispersion, 240

Homologous line pairs, 283

Hüfner spectrophotometer, 378

Hurter and Driffeld curves, 424

Huygens eyepiece, 83

Huygens principle, 156, 210

Hydrogen source for absorption, 401

Hyperfine structure, 145, 295

Hyperfocal distance, 57

Hyperopia, 77

I

Ideal optical systems, 10, 22

Illuminance, 62

measurement of, 346

of images, 66

Images, real and virtual, 10

Image space, 11

Independence of light rays, 3

Index, ellipsoid, 205

of absorption, 237, 244

of refraction, absolute, 2, 109

measurement of, by critical angle, 362

by minimum deviation, 357

relative, 2, 109

Index surface, 208
 Infrared plates, 426
 Infrared spectrometer, 380
 Infrared thermal radiation, 263
 Intensity of source, 62
 Interference, constructive, 116
 destructive, 116
 in films and plates, 124
 with large path difference, 133
 with polarized light, 215
 Interferometer, Fabry-Perot, 147
 Jamin, 146
 Michelson, 134, 368
 Michelson stellar, 121
 Rayleigh gas, 123
 Internal standard lines, 399
 International angstrom unit, 146
 Inverse square law, 63
 Iron-arc standard lines, 394
 Isochromatic lines, 218
 Isoclinics, 217
 Isotropic media, 109

J

Jacobs, optical manufacture and assembly, 89, 444
 Jamin, ellipticity at polarizing angle, 254
 Jamin interferometer, 146
 Jenkins and White, interference spectroscopes, 153
 physical optics text, 441
 Johansson gauge blocks, 128, 154, 258

K

Katoptric systems, 17
 Kellner eyepiece, 84
 Kellstrom, x-ray diffraction, 186
 Kerr cell, 321
 Kerr electro-optic effect, 319
 Kerr magnetic effect, 313
 Ketteler's formulas for metals, 246
 King, microwave absorption, 295
 Kingslake, photographic objectives, 48
 Kirchhoff, diffraction theory, 173, 183
 radiation law, 259

Kirchhoff, solution of the wave equation, 156, 173, 183
 Kramers, x-ray dispersion, 238
 K series, 302
 Kundt, thin metallic films, 247

L

Lagrange's law, 23
 Lambert, cosine law, 64
 unit of luminance, 63
 Lambert-Beer absorption law, 349
 Landé splitting factor, 317
 Larsson, x-ray dispersion, 240
 x-ray spectrometer adjustment, 413
 Laurent plate, 231, 351
 Lieber and Lieber, relativity, 140
 Light, flux, 61
 nature of, 106, 323
 vector, 107, 203
 Light sources, a-c arc, 285
 color temperatures, 269
 Fuessner spark, 286
 r-f discharge, 376
 hydrogen arc, 401
 Lindh, x-ray spectra, 310
 Linear interpolation in spectra, 396
 Linear magnification, 14
 Line spectra, 270
 Lippich half-shade, 230, 351
 Littrow quartz spectrograph, 394
 Localization of fringes, 128, 137
 Longitudinal magnification, 18
 Lorentz splitting, unit, 317
 L series, 302, 305
 Lumen, 61
 Luminance, 63
 Luminosity, 327
 Lummer-Brodhun photometer, 344
 Lummer-Gehrcke plate, 153
 Lundquist, x-ray spectra, 310
 Lux, 63
 Lyman series, 272

M

Macbeth illuminometer, 346
 Mach interferometer, 146
 Mack and Martin, photography, 444

Magnetic double refraction, 313
 Magneto- and electro-optics, 312
 Magnification by spectacle lenses, 80
 Magnifications related, 18
 Magnifying glass, 81
 Magnifying powers of telescopic systems, 32
 Malus, cosine squared law, 193
 Mass absorption coefficient, 302
 Maxwell's equations, 194
 for conducting media, 242
 Mean spherical candle power, 62
 Meggers, standard wave length, 146
 Mercury, arc spectrum, 279, 433
 energy levels, 278
 198 isotope, 146
 Zeeman effect, 314
 Mesons, 324
 Metallic media, 242
 measurement of indices, 407
 Metals and insulators compared, 253
 Meyer, diffraction grating, 171
 Michelson, ether drift experiment, 138
 star diameters, 121
 studies in optics, 112, 142
 Michelson echelon grating, 153
 Michelson interferometer, 134
 adjustment of, 368
 white light fringes, 371
 Microphotometer, 399
 Microscope, 85
 Microwave absorption, 294
 Microwave diffraction, 185
 Miller, chemical deposition of silver, 443
 ether drift experiment, 138
 Miller indices, 187
 by powder method, 418
 Minimum deviation, 92
 measurement of, 359
 Mirrors, aspheric, 4
 reflection from, 7, 23
 Molecular beams, 296
 Monk, evaporation of metals, 443
 mounting of cross-hairs, 443
 Moseley's law, 300
 Mott, latent image, 422
 Multiplicity in line spectra, 276

Myopia, 75

N

Near point, 75
 Neutrons, 324
 Newtonian telescope, 90
 Newton's color scale, 127
 Newton's rings, 131
 Nicol prism, 213
 Nodal points, 17
 Nodal slide, 338
 Non-diagram lines, 307
 Normal brightness, 68
 Normal magnification, 69
 Normal reflection by metals, 247
 Normal slit width, 97
 Nuclear magnetic moment, 297
 Numerical aperture, 66, 86

O

Oblique reflection, by insulators, 199
 by metals, 250
 Oculars, 83
 Opacity, 380
 Ophthalmic lenses, 75
 Optical activity, 227
 Optical bench, 338
 Optical constants of metals, 243
 Optical flats, 129, 353
 Optical instruments, 73-100
 Optical path, 3
 Optic axis, 207
 Object space, 11
 Orthoscopic lens, 43
 Ostwald branch, 79

P

Page, linear oscillator, 265
 Paraxial rays, 22
 Paschen-Back effect, 319
 Paschen series, 272
 Pease and Pearson, speed of light, 112
 Pellin-Broca prism, 94
 Pencil of rays, 1
 Period, 108
 Persistent lines, 281
 Perspective, binocular, 101

- Perspective, monocular, 58
 Petzval, condition for flatness, 49
 Petzval surface, 47
 Pfund, channelled spectrum, 389
 Pfund arc, 394
 Pfund refractometer, 362
 Pfund series, 272
 Phase, 108
 Phase shift, beyond critical angle, 256
 by metals, 248
 on reflection, 125, 201
 Phot, 63
 Photoelasticity, 217, 218, 225
 Photoelectric cells, 322
 conduction, 323
 emission, 322
 Photoelectric effect, 322
 Photographic process, 421-429
 densitometry, 284, 424
 development, 423
 fixation, 424
 granularity, 427
 latent image, 423
 latitude, 426
 objectives, 36, 48
 photometry, 421, 429
 reciprocity law, 428
 sensitometry, 425
 Photometry, 61, 344
 Photon, 323
 Photovoltaic cells, 323
 Physical limits to image quality, 36
 Physical optics, bibliography, 441
 Planck's distribution law, 265
 Plane, of incidence, 2
 of polarization, 107
 Plane grating, 168, 365
 Plane wave, equation, 107
 Plate calibration, 284
 Polarimetry, 230
 experiment, 349
 Polaroscope, circular, 221
 plane, 215
 Polarization by scattering, 333
 Polarization ovaloid, 204
 Polarized light, 193
 Polarizing angle, 201
 Polaroid, 203, 212
 Porro prisms, 89
 Power of a lens, 29
 Poynting's vector, 198
 Prentice's rule, 79
 Presbyopia, 75
 Primary colors, 325
 Primary focus, 46
 Principal angle of incidence, 252
 Principal axes, 194
 Principal azimuth, 252
 Principal foci, 13
 Principal maxima in diffraction, 169
 Principal planes, of crystal, 210
 of lens, 14, 29
 Principal section, 214
 Principal vibration directions, 212
 Prism diopters, 79
 Prisms, Cornu, 392
 Fresnel, 228
 ophthalmic, 79
 Pellin-Broca, 94
 Porro, 89
 quartz, 392
 rock-salt, 380
 spectrometer, 91
 Prism spectrograph, 91
 Prism spectrometer, 91
 adjustment of, 358
 Prism spectroscope, 91
 Purity factor, 98
- Q
- Qualitative spectrochemical analysis,
 281, 388
 Quantitative spectrochemical analysis,
 283, 398
 plate calibration, 429
 Quantum mechanics, 279
 Quantum numbers, 276
 Quantum theory, 106
 band spectra, 286-294
 dispersion, 238
 line spectra, 270-286
 photoelectricity, 322
 thermal radiation, 267
 Zeeman effect, 317
 Quarter-wave plate, 219

Quartz spectrograph, 391

R

Rabi, molecular beam experiments, 297

Radiant flux, 198

Radiation and spectra, 259-293
bibliography, 442

Radius of stereoscopic vision, 101

Raman effect, 335

Ramsden eyepiece, 84

Ratio, of focal lengths, 24
of wave lengths, 152, 370

Rayleigh gas interferometer, 123

Rayleigh-Jeans formula, 267

Rayleigh scattering, 331

Rays, 1, 207

Ray velocity surface, 207

Rectilinear lens, 44

Rectilinear propagation, 1

Reflecting telescopes, 90

Reflection, and conductivity, 249

by metals, 247

normal, 248

oblique, 250

electromagnetic theory, 199

laws of, 2, 199

of polarized light, 404

total, 254

Reflection coefficient, 8

Reflectivity, diffuse, table, 434
specular, table, 434

Refraction, at single surface, 22

at spherical surface, 20

electromagnetic theory, 199

electron theory, 232

laws of, 2, 6, 200

of x-rays, 240

quantum theory, 238

Refractive indices, tables, 434, 435

Refractivity, 233

Refractometer, parallel plate, 362

Relative aperture, 54, 68

Relativity theory, 140

Resolving power, of Fabry-Perot in-
terferometer, 150

of grating, 171, 367

of instruments, 163

Resolving power, of microscope, 86
of spectroscope, 95, 360

Retina, 74

Reversibility principle, 7

Rochon prism, 226

Roemer, speed of light, 114

Rosa and Dorsey, speed of light, 112

Rotational bands, 287

Rotatory dispersion, 231

Rowland circle, 385

Rydberg constant, 275

Rydberg series formula, 275

S

Sagittal focus, 46

Sawyer, spectroscopy, 399, 400

Scattering of light, 331

Schmidt telescope, 90

Schrodinger, wave mechanics, 279

Schuster, elliptically polarized light,
407

spectral purity, 97

Screening constant, 300

Secondary beta rays, 323

Sector photometry, 401

intermittency effect, 403, 428

Seidel's aberration theory, 39

Selection rules, 277, 280

Selective thermal radiation, 267

Sensitive lines, 281

Sensitizing for ultraviolet, 426

Series, spectral, 271-279

Shape of fringes, 138

Shillaber, microscopy, 87

Siegbahn transposition method, 413

Siegbahn vacuum spectrometer, 416

Sign convention, for ideal systems, 13,
17

for mirrors, 23

for single refracting surfaces, 20

Simple lens, 27

Sine condition, 36

Single-slit diffraction, 157, 181

Sky, color of, 333

Slope angle, 16

Snell's law, from electromagnetic the-
ory, 200

from Fermat's principle, 7

Soliel-Babinet compensator, 221
 Sommerfeld, fine structure, 276
 Sonnar lens, astigmatism, 48
 Sources for spectrochemical analysis,
 281, 286
 Sources of light, a-c arc, 285
 color temperatures, 269
 Fuessner spark, 286
 r-f discharge, 376
 hydrogen arc, 401
 Specific intensity, 259
 Specific refractivity, 233
 Specific rotatory power, 227, 352
 Spectacle lenses, 75
 Spectral series, 271
 Spectrochemical analysis, 270, 281
 Spectrographs, 91-98
 grating, 383
 Littrow, 394
 quartz, 392
 Spectrometer adjustment, 358
 Spectrophotometry, 377
 Spectroscope calibration by interfer-
 ence, 389
 Spectroscopic terms, 274
 Spectrum line curvature, 95
 Speed indices for plates and films, 425
 Spherical aberration, 39
 measurement of, 341
 Spherical mirror, 23
 Sphero-cylinders, 78
 Spherometer, 30
 Speed, of electromagnetic waves, 112
 of light, 112
 Standard illuminants, 325
 Star, contrast, 71
 diameter of, 123
 Stefan-Boltzmann law, 266
 Stereoscopic camera, 101
 Stereoscopic viewer, 102
 Stereoscopy, 101-105
 Stops, 53
 Straight-edge diffraction, 180
 Strong, coating of mirrors, 443
 filters, 443
 sources of light, 444
 Sturm's conoid, 45
 Subjective brightness, 68

Surface testing by interference, 129,
 352

T

Tangential focus, 46
 Taylor, mirrors by sputtering, 443
 Telescopes, 87
 Telescopic systems, 31
 Terms, spectroscopic, 275
 Terrestrial telescope, 89
 Thermal radiation, 259-269
 Thick lens, 28
 Thickness by interferometer, 371
 Third order aberrations, 38
 Thomson, scattering formula, 332
 Tolansky, interferometry, 153
 light sources, 444
 Toothed disk, speed of light by, 113
 Total reflection, 254
 Transmissivity, 379
 measurement of, 401
 Transposition, 78
 Transverse magnification, 14
 Transverse waves, 197
 Trichromatic colorimetry, 325
 Trigonometrical ray tracing, 20
 Tscherning ellipse, 79
 Tyndall effect, 331

U

Ultraviolet, absorption, 391, 401
 emission, 391
 Uniaxial crystals, 203
 Unpolarized light, 108
 Urey, hydrogen tube, 401, 444

V

Valasek, optical constants of metals,
 247
 x-ray spectra, 310
 Van Vleck, absorption of microwaves,
 295
 Vectographs, 104
 Vertex focal length, 30
 Vertex power, 30, 77
 Vibration-rotation bands, 289

Vignetting, 56
 Virtual focus, 10
 Visibility, of interference fringes, 140, 369
 Visibility function for eye, 61, 436
 Visible spectrum, 108
 Von Seidel formulas, 39

W

Wadsworth mirror, 94
 Wave front, 108
 Wave lengths, 108
 M.I.T. tables, 282
 measurement of, by grating, 365
 by interpolation, 393
 mercury arc, 357, 432, 433
 ratio by interferometer, 153, 374
 standard, 146
 useful values, 432
 Wave mechanics, 279
 Waves, addition of, 116
 Wave theory, 107
 Wave velocity surface, 207
 Weak field Zeeman effect, 318
 Webb, latent image, 422
 Weston foot-candle meter, 346
 Wetterblad, x-ray spectra, 310
 Widths of spectral lines, 142
 Wien, displacement law, 262
 distribution formula, 267
 Williams, interference, 153
 Wilsey, optical constants of gold, 246

Wollaston branch, 79
 Wollaston prism, 226
 Wood, magneto-optics, 312
 ruling of gratings, 171
 thickness by interference, 371

X

X-ray, absorption coefficient, 302
 absorption edges, 301
 circuit for operating tube, 412
 diffraction by crystals, 186, 417
 dispersion of, 239
 energy levels, 302
 general radiation, 310
 measurement of wave lengths, 409, 415
 satellite lines, 307
 scattering, 302, 332
 spectra, 299
 spectroscopy, 190, 409
 tube, 411
 unit of wave length, 410

Y

Young, interference experiment, 115

Z

Zeeman effect, 314
 Zehnder interferometer, 146
 Zeiss Sonnar lens, 48
 Zone plate, 183

DATE OF ISSUE

This book must be returned within 3/7/14 days of its issue. A fine of ONE ANNA per day will be charged if the book is overdue.

--	--	--	--	--	--

